

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:45:20 ; Search time 36.6935 Seconds

(Without alignments)
3043.156 Million cell updates/sec

Title: US-09-927-315-9

Perfect score: 4443
Sequence: 1 MGPRACTICLFLWLWLAE.....ERNTPAVFNSMIGGYTWRD 838

Scoring table: BIOSOM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3514	79.1	669	21	AAV77558 Human GPCR-B4 poly
2	3203.5	72.1	843	21	AAV77556 Rat GPCR-B4 poly
3	3151.5	70.9	843	21	AAV77557 Mouse GPCR-B4 poly
4	1454	32.7	840	21	AAV45021 Rat sensory transd
5	1453	32.7	840	21	AAV45028 Rat sensory transd
6	1452	32.7	840	21	AAV45027 Rat sensory transd
7	1451	32.7	840	21	AAV45026 Rat sensory transd
8	1434	32.3	842	21	AAV45022 Rat sensory transd
9	1354	30.5	841	22	AAE10372 Mouse sensory tran
10	1354	30.5	841	22	AAE11969 Human novel G-prot

11	1354	30.5	841	23	ABW77319 Human G-protein co
12	1270	28.6	777	21	AAV45023 Human sensory tran
13	1267.5	28.5	1138	23	AAO21502 Human novel G-prot
14	1260.5	28.4	763	22	AAE11970 Chicken calcium-se
15	1107	24.9	1059	22	AAU00508 Dogfish shark kidn
16	1100	24.8	1026	18	AAW32059 Parathyroid calculi
17	1094	24.6	1078	19	AAW11869 Human parathyroid
18	1094	24.6	1078	19	AAW54846 Human parathyroid
19	1094	24.6	1078	19	AAW38274 Human parathyroid
20	1094	24.6	1078	20	AAV28840 Human calcium rece
21	1094	24.6	1078	20	AAV41780 Human parathyroid
22	1094	24.6	1078	20	AAW69565 Human parathyroid
23	1094	24.6	1078	21	AAV51827 Human calcium rece
24	1094	24.6	1078	21	AAV70325 Human wild type ca
25	1094	24.6	1078	22	AAV47391 Protein encoded by
26	1094	24.6	1078	23	AAV47822 HuCAR4.0. Homo sa
27	1093	24.6	1078	22	AAU02195 Cynomolgous monkey
28	1092.5	24.6	1027	23	ABW78761 Dogfish shark kidn
29	1092.5	24.6	1027	23	AAU76004 Shark kidney calculi
30	1088.5	24.5	1079	19	AAW54847 Rat kidney calcium
31	1088.5	24.5	1079	19	AAW38275 Rat kidney cell ca
32	1088.5	24.5	1079	20	AAV41781 Rat parathyroid ca
33	1088.5	24.5	1079	20	AAW34928 Rat kidney extrace
34	1088.5	24.5	1079	20	AAW69566 Rat parathyroid ca
35	1088.5	24.5	1079	21	AAV51828 Rat calcium recept
36	1088.5	24.5	1079	23	AAW47823 RakCAR3A. Rattus
37	1087.5	24.5	1085	19	AAW54844 Bovine parathyroid
38	1087.5	24.5	1085	19	AAW38272 Bovine parathyroid
39	1087.5	24.5	1085	20	AAV41778 Bovine parathyroid
40	1087.5	24.5	1085	20	AAW69563 Bovine parathyroid
41	1087.5	24.5	1085	21	AAV51825 Bovine calcium rec
42	1087.5	24.5	1085	23	AAW47820 BOPCAR1. Bovine s
43	1079	24.3	1088	17	AAW11888 Parathyroid calculi
44	1079	24.3	1088	19	AAW54845 Human parathyroid
45	1079	24.3	1088	20	AAV41779 Human parathyroid

ALIGNMENTS

RESULT 1	AAV77558 standard; Protein: 669 AA.
XX	AAV77558
XX	AAV77558;
AC	08-MAY-2000 (first entry)
XX	Human GPCR-B4 polypeptide.
DE	Sensory transduction G-protein coupled receptor; GPCR; GPCR-B4; human;
XX	taste transduction pathway; taste receptor; foliate; fungiform; food;
KW	circumvallate; taste signaling; pharmaceutical.
KW	
XX	
OS	Homo sapiens.
XX	
PN	WO200006593-A1.
XX	
PD	10-FEB-2000.
XX	
PF	27-JUL-1999; 99WO-US17104.
XX	
PR	28-JUL-1998; 98US-0095464.
XX	PR 17-DEC-1998; 98US-0112747.
XX	(REGC) UNIV CALIFORNIA.
XX	Zucker CS, Adler JE, Lindemeyer J;
XX	WPI; 2000-195257/17.
XX	N-PSDB; AAZ58965.
PT	New isolated sensory transduction G-protein coupled receptor, useful

Db 121 DDLPLPLKDYSGYMPHYVAVIGPDNSESATVSNILSHFLIPQITYSALISDKLRKRF 180
 QY 178 PALLRTTPSADHVEAMVQLMHFRRNMATIVYSSDTYGRDNGQLGGEVVAR-RDICIATF 236
 Db 181 PSMLETRVPATHHIEAMVQLMVHFQMNMTIVLVSDDDYGERNSHLSQRLTFTSDICIAF 240
 QY 237 QETLPTLPONQMTSEERQRLVTIYDKLQOSTARVVVYSPDLTYHFENEYLRONFGA 296
 Db 241 QEVLPVPESQYMRSEEQQLNDIKLKRISARVVVYSPDLTYHFENEYLRONFGA 300
 QY 297 VVIASESWAIDVPLNHLNTELGHLGTFGLTITQSVPIPGSEPREMGPOAGPPLSTQS 356
 Db 301 VVIASESWAIDVPLNHLNTELGHLGTFGLTITQSVPIPGSEPREMGPOAGPPLSTQS 360
 QY 357 YTCNOCDCACNMATSEFNILSLGSRVYVYSAVYAAHHLHLGCGDKCTCKRKYV 416
 Db 361 YTCNOCDCACNMATSEFNILSLGSRVYVYSAVYAAHHLHLGCGDKCTCKRKYV 420
 QY 417 PMQLLEIKVNFLLDHOIFFDPOGVALHLEIYQWMDRSONPFQSVASYYPLOROLK 476
 Db 421 PMQLLEIKVNFLLDHOIFFDPOGVALHLEIYQWMDRSONPFQSVASYYPLOROLK 480
 QY 477 NIQDISMHTVNTTIPMSCKRCQSGQKKRPVGIHVCECICIDCLPGTFNLHTEDEYEQ 536
 Db 481 YINNVSWYTPNNTVPIVMSCKRCQSGQKKRPVGIHVCECICIDCLPGTFNLHTEDEYEQ 540
 QY 537 ACPNEMMSQSETEKROLVLEWHEARTIYVALLAAGFISTALIVIFRRHQTPIV 596
 Db 541 SCQGSWSTKNDITCFQRPRITLWHEVPTIVALLAAGFISTALIVIFRRHQTPIV 600
 QY 597 RSAGGPMCFMLTLNLAIVAVVYVYVGPVKVSTCLCRLAFLPCEFTICISCIATVASFQIV 656
 Db 601 RSAGGPMCFMLTLNLAIVAVVYVYVGPVKVSTCLCRLAFLPCEFTICISCIATVASFQIV 660
 QY 657 CAFKASRPRAVSTVWYQGPVYSMAFTVTKMVIIVIGLARPOSH--RTDDDPKRI 714
 Db 661 CYFKARRLPSAYSEFMRHGPYVAFITAKVALVGNMIA--TINPIGRDTPDPPI 719
 QY 715 TVSCNPNRNSLLENTSLDILLVYGFSAVGMKELPNTYNAKRTITSMTEFTSSVS 774
 Db 720 MTLSCHPNRNGLTNTSMDLLSVLGFSAVGMKELPNTYNAKRTITSMTEFTSSVS 779
 QY 775 LCTFMSASVGLVITVDLVYVNLNLATLSLGFYGPCKYMLTFYPERNTPAYNSMIOGYT 834
 Db 780 LCTFMSASVGLVITVDLVYVNLNLATLSLGFYGPCKYMLTFYPERNTPAYNSMIOGYT 839
 QY 835 MRR 837
 Db 840 MRR 842

RESULT 3

AAV77557 standard; Protein; 843 AA.

AAV77557;

08-MAY-2000 (first entry)

Mouse GPCR-B4 polypeptide.

KW Sensory transduction G-protein coupled receptor; GPCR; GPCR-B4; mouse;
 taste transduction pathway; taste receptor; foliate; fungiform; food;
 circumvallate; taste signaling; pharmaceutical.

Mus sp.

WO200006593-A1.

10-FEB-2000.

27-JUL-1999; 99WO-US17104.

XX

PR 28-JUL-1998; 98US-0095464.
 PR 17-DEC-1998; 98US-0112747.
 PA (REGC) UNIV CALIFORNIA.
 PI Zuker CS, Adler JE, Lindemeyer J;
 DR WPI: 2000-195257/17.
 DR N-PSDB; AAZ38964.
 PR New isolated sensory transduction G-protein coupled receptor, useful
 PT for developing products for use in studying and modulating the taste
 PT transduction pathway and for generating taste topographic maps -
 PS Claim 22; Page 69-70; 76pp; English.
 CC The invention provides nucleic acids encoding rat, mouse and human
 CC sensory transduction G-protein coupled receptor (GPCR) polypeptides. The
 CC GPCR polypeptides are components of the taste transduction pathway. The
 CC nucleic acids can be used to identify taste cells and as tools for the
 CC generation of taste topographic maps that elucidate the relationship
 CC between the taste cells of the tongue and taste sensory neurons leading
 CC to taste centers in the brain. GPCR-B4 is useful as a nucleic acid probe
 CC for identifying subpopulations of taste receptor cells such as foliate,
 CC fungiform, and circumvallate taste receptor cells. The polypeptides can
 CC be used for identifying compounds that modulate sensory signaling in
 CC sensory cells. Such modulators of taste transduction are useful for
 CC pharmacological and genetic modulation of taste signaling pathways. These
 CC modulatory compounds can then be used in the food and pharmaceutical
 CC industries to customize taste. The present sequence represents a
 CC mouse GPCR-B4 polypeptide.
 SQ Sequence 843 AA.
 Query Match 70.9%; Score 3151.5; DB 21; Length 843;
 Best Local Similarity 68.9%; Pred. No. 7.6e-295;
 Matches 581; Conservative 113; Mismatches 142; Indels 7; Gaps 4;
 QY 1 MGPRAKTIQSLFFLLVLAEP---AENSDFYLPQDYLGLSLHANNMGYHIANLOYP 57
 Db 1 MGPRAKTIQSLFFLLVLAEP---AENSDFYLPQDYLGLSLHANNMGYHIANLOYP 60
 QY 58 MCKEYEVKIGYNLQAMAFVEEINNDSSLPGVLYGEIYDVCYISNNQVPLYLELAH 117
 Db 61 KNEINEMKATGLINLQAMAFVEEINNDSSLPGVLYGEIYDVCYISNNQVPLYLELAH 120
 QY 118 EDNLLPILQEDYSNYISRVVAVIGPDNSESATVSNILSHFLIPQITYSALISDKLRKRF 177
 Db 121 IDDLPLPLKDYSGYMPHYVAVIGPDNSESATVSNILSHFLIPQITYSALISDKLRKRF 180
 QY 178 PALLRTTPSADHVEAMVQLMHFRRNMATIVYSSDTYGRDNGQLGGEVVAR-DICIAF 236
 Db 181 PSMLETRVPATHHIEAMVQLMVHFQMNMTIVLVSDDDYGERNSHLSQRLTFTSDICIAF 240
 QY 237 QETLPTLPONQMTSEERQRLVTIYDKLQOSTARVVVYSPDLTYHFENEYLRONFGA 296
 Db 241 QEVLPVPESQYMRSEEQQLNDIKLKRISARVVVYSPDLTYHFENEYLRONFGA 300
 QY 297 VVIASESWAIDVPLNHLNTELGHLGTFGLTITQSVPIPGSEPREMGPOAGPPLSTQS 356
 Db 301 VVIASESWAIDVPLNHLNTELGHLGTFGLTITQSVPIPGSEPREMGPOAGPPLSTQS 360
 QY 357 YTCNOCDCACNMATSEFNILSLGSRVYVYSAVYAAHHLHLGCGDKCTCKRKYV 416
 Db 361 YTCNOCDCACNMATSEFNILSLGSRVYVYSAVYAAHHLHLGCGDKCTCKRKYV 420
 QY 417 PMQLLEIKVNFLLDHOIFFDPOGVALHLEIYQWMDRSONPFQSVASYYPLOROLK 476
 Db 421 PMQLLEIKVNFLLDHOIFFDPOGVALHLEIYQWMDRSONPFQSVASYYPLOROLK 480
 QY 477 NIQDISMHTVNTTIPMSCKRCQSGQKKRPVGIHVCECICIDCLPGTFNLHTEDEYEQ 536
 Db 481 YINNVSWYTPNNTVPIVMSCKRCQSGQKKRPVGIHVCECICIDCLPGTFNLHTEDEYEQ 540

RESULT 5
AAV45028 standard; Protein; 840 AA.
ID AAV45028 standard; Protein; 840 AA.
XX
AC AAV45028;
XX
DT 31-MAY-2000 (first entry)
XX
DE Rat sensory transduction G-protein coupled receptor-B3 variant #3.
XX
KM Rat; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
KM sensory cell; taste receptor cell; screen; taste modulator;
KM pharmaceutical; food industry; taste topographic map; tongue; variant.
XX
OS Rattus sp.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 90 /note= "Wild type Ala substituted by Gly"
XX
XX WO200006592-A1.
XX
XX 10-FEB-2000.
XX
XX 27-JUL-1999; 99WO-US17099.
XX
XX 28-JUL-1998; 98US-0094465.
XX
XX (REGC) UNIV CALIFORNIA.
XX (USSS) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Zuker CS, Adler JE, Lindemeier J, Ryba N, Hoon M;
XX WPI; 2000-205451/18.
XX
XX New isolated sensory transduction G-protein coupled receptor, useful
XX for developing products for use in studying and modulating the taste
XX transduction pathway .
XX
XX Disclosure; Page -: 83pp; English.
XX
XX The present sequence is a polymorphic variant #3 of rat taste cell
XX specific G-protein coupled receptor, GPCR-B3 which is involved in
XX sensory transduction. GPCR-B3 is specifically expressed in foliate
XX and fungiform cells, with lower expression in circumvallate taste
XX receptor cells of the tongue. The protein has an extracellular domain,
XX seven transmembrane domains and an intracellular domain.
XX The GPCR-B3 sequence is used to screen compounds that modulate sensory
XX signalling in taste cells, especially taste modulators useful in
XX pharmaceutical and food industries to customise taste. The sequence
XX can also be used as probe for identifying taste cells and
XX subsets of taste receptor cells such as foliate, fungiform and
XX circumvallate. Such probes are also useful to generate taste
XX topographic maps that elucidate the relationship between the taste
XX cells of the tongue and sensory neurons leading to taste centres
XX in the brain.
XX Note: The present sequence is not given in the specification but is
XX derived from rat GPCR-B3 sequence shown in page 75 (AAV45021).
XX
XX Sequence 840 AA:
XX
Query Match 32.7%; Score 1453; DB 21; Length 840;
Best Local Similarity 39.5%; Pred. No. 9,1e-131;
Matches 325; Conservative 122; Mismatches 347; Indels 28; Gaps 11;
OY 27 FYLPDGYILGGLFSLHANNKGIHLNLFQVPMG-KEYEKVIGYNLMQMRFAVEINND 85
DB 31 FSLPDGFLAGLFLSHGDLQVHRPL--VISCDRPDSNGGHHLEFQMRFTVEINNS 88
OY 86 SLLPGLVLLGVEIVDVCYISNNVQVLYFLAHE-DNLAPIQDYSNYISRYVAVIGPDNS 144

DB 89 SGLLPNITLIGVLYDVCCSSANVYATIRIALAQGRHIEIQDLANHSSKVAFIGPDNT 148
OY 145 ESMVTVANFLSLFLLPQITYSASIDELNDKVFALLRTTSPADHHEAMVOLMHEFRNN 204
DB 149 DHAVTTAALIGFLMPLVSYEASSVLSAKRFPSFLRTVPDRHQVEVMQVLLQSPGV 208
OY 205 WIIIVSSDPTVGDRNGQLGGERVARRDICIAQFELPLTPQPNMTSEERQVLYIVDKL 264
DB 209 WISLIGSDYDGLQVQALAEELAVPRGICVAKDIPV-----SARVGDPPHQMOMHL 262
OY 265 OOSTARVVVFSPDLTLVHFENEVRONFTGAWMVASSEMAIDPVLHTELHGATFLG 324
DB 263 AQARTTVVVFNSRLIARVFRSVLANITGVWASEDMAISTYITSTYIGIGIYVIG 322
OY 325 ITIGSVPIGSEFSEFREMWDQAGPPLSRTSOSTYC--NOECDNCLNATLSEFTIILRSGE 382
DB 323 VAVQOROVPGLEKEFESYBRAVTAAPSACPEGSMCSTNQLCRECHTFTRNNPTLGAFSM 382
OY 363 RYVYSVYSAVYVAHALSHLQCDKSTCKRYVYPMOLLEIKVNFLLDHQIFPDQ 442
DB 363 SAAYRYEAVYVAHAGLHQLGCTSEICSRGVYPMOLQIYKVFLLHENTVAFDNG 442
OY 443 DVALLHLEIVOMQDRSONPQSV--ASYYPLOQLKNIDISMTVNNTPMMSCKRQ 500
DB 443 DPLGYDDIAMDNDMPENTFELIGSASLSPVHLDI-NKTIOMHKKNNQVPVSCITDCL 501
OY 501 SGQKKRPVGIHVCFECIDCLPGTFLNTHEDYEQACQPNMWSYQSTSCFRQLVLE 560
DB 502 AGHHRVAVSHHCCEVCCEAGTEFLNMSE-LHICQPGTEEMAKRESYTCPRVFEFLA 560
OY 561 WHEAPTIIVALLAAGFLSTALIV-----TFMHHPOPIYRSAGGPGCFMLLLVLA 614
DB 561 WHE--PISLVLTAA---NTLLLLLVGTAGLEAFNHFTPVVRSAGRCFTMLSLVAG 614
OY 615 YWVVPVYVGPVSTCLRQALFPLFTICISCIARFQIACAKMASREFRAYSVMW 674
DB 615 SCSFYSPFGEPTVPACILRQPLFSIGFAIFLSCLIRSQVLIIRKFSKVTFRMAQ 674
OY 675 YQPVYSAFTVLMKAVIVIGMLARPOSHPTDDPKITIVSCNPNRSLNLENTSLD 734
DB 675 NHGAGLFVIVSSTVHLILCLTWLVMTPTREYORFPHLVILCEVNSVGFLLAFTHN 734
OY 735 ILLSVGFSFAMGKELPYNENAKFTILSMFYTSSVSLCTFMSAVSGLVLTVDLLV 794
DB 735 ILLSTIVCSYLGELENENAKCVTFSLLNFVMAFTMASTYIGSTLPAVNVLA 794
OY 795 TVLNLAIISLGYGPKCYMILFYPERNTPAYENSMIOGYTR 836
DB 795 GLTTLISGFSGYFLPKCYVILCRPRLNTEHQAASIDQYTR 836
RESULT 6
AAV45026 standard; Protein; 840 AA.
ID AAV45026 standard; Protein; 840 AA.
XX
AC AAV45026;
XX
DT 31-MAY-2000 (first entry)
XX
XX Rat sensory transduction G-protein coupled receptor-B3 variant #1.
XX
XX Rat; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
XX sensory cell; taste receptor cell; screen; taste modulator;
XX pharmaceutical; food industry; taste topographic map; tongue; variant.
XX
XX Rattus sp.
XX Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 33 /note= "Wild type Leu substituted by Ile"
XX

QY 483 HTVANNITPMSCSKRCOSGOKKRPVGIHVCCFECIDCLPGFELNHTDEDECOACPNNE 542
 DB 485 WHGKNOVPKSYCSDCLBEGHQRVVTGHHCCFECVPCGAGFLNKS-DLNRQPCGKEE 543
 QY 543 WSYOSETCFKROLVLEHMEAPITAVALLAAGFLSTLALIVFWRHQPTIVSAGP 602
 DB 544 WAPESGQCFPRVTVFALREHTSMVLLAANTLLLLLLGTLGFLAMHLDPPVRSAGR 603
 QY 603 MCFMLTLILVAVVVPVYVGPVKSTCLCQALPPLCFTTICISIAVRSFOIYCAFMA 662
 DB 604 ICFMLGSLAAGSGSLVGFEGFPTRACLLRQALFALFTFLSCILTVRSFOLLITIFRS 663
 QY 663 SRFPRAYSYWRVYGPVYSMAFIVLAKVIYIGMLARPOSHPRTPDDPKITIVSCPN 722
 DB 664 TKVPTFYHAWQNHGAGLEFVMISSAQLLCLTLVWVTPPAREYORFPLVMECTET 723
 QY 723 YRNSL-----LFTNSLDLLSVGFSFAYMGKELPTYNNAKFTTSMTEYFSSVSLC 776
 DB 724 --NSIGFLIAFLYNG----LTSISAFACSYLCKDLPENNNEAKCVTSLFNFVSMIAFE 777
 QY 777 TFMASVSGVLVTVIDLLTVNLALISIGYFEPKCYMILFYERTPRAYFNMIQYIMR 836
 DB 778 TTAASYDGYLTPAAMMAGLSSLSGFGYFLPKCYVILCRPDLNSTERHAFASIDYTRR 837
 RESULT 10
 AAE11969
 ID AAE11969 standard; Protein: 841 AA.
 AC AAE11969;
 XX
 DT 18-DEC-2001 (first entry)
 DE Human novel G-protein coupled receptor (NGPCR) protein #1.
 XX
 KW Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
 KW antisense-therapy; signal transduction; behavioural disorder; obesity;
 KW heartbeat rate; inflammation; immune disorder; diabetes; cancer;
 KW coronary disease.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 107
 FT /note= "Encoded by TTT"
 FT Misc-difference 372
 FT /note= "Encoded by RCA"
 FT
 XX WO200172842-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 28-MAR-2001; 2001WO-US09996.
 XX
 PR 28-MAR-2000; 2000US-192978P.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abulin A;
 PI Zambrowicz B, Sands AT;
 DR WPI: 2001-616474/71.
 XX
 N-PSDB: AAD19501.
 XX
 PT Novel isolated polynucleotides encoding human G protein coupled
 PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
 PT expressed NGPCRs for diagnosis of disease, and as probes or primers
 XX
 PS Claim 3; Page 72-74; 80pp; English.
 CC The present sequence is human novel G-protein coupled receptor
 CC (NGPCR) protein. NGPCRs are transmembrane proteins that span

CC the cellular membrane and are involved in signal transduction after
 CC ligand binding. The NGPCR polynucleotide sequences, are useful
 CC in diagnosis and treatment of a disease involving NGPCR, for detecting
 CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NGPCR in
 CC the body or abnormalities in the signal transduction pathway mediated by
 CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
 CC trial monitoring and/or the treatment of physiological (heartbeat rate)
 CC or behavioural disorders. NGPCR is useful for identifying compounds
 CC useful in the therapeutic treatment of obesity, inflammation, immune
 CC disorders, diabetes, heart and coronary disease, metabolic disorders and
 CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
 CC genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests.
 CC
 XX Sequence 841 AA;
 SO
 Query Match 30.5%; Score 1354; DB 22; Length 841;
 Best Local Similarity 37.4%; Pred. No. 3.3e-121;
 Matches 314; Conservative 131; Mismatches 355; Indels 40; Gaps 16;
 QY 16 WYLA-EPAENS-DEYLPDDYLGLFSLHANKGIVHLNQLQVPMC-KEYEYKIVGMIM 72
 DB 19 WAFACHSTESSPDFTLPDDYLAGLFLHSGCLQYRHRP--EVTLCDRSCSPNEGHYHLF 76
 QY 73 QAMRAVEEINNDSSLRLGVLGELYDYCYISNNVQVLYFLA---HEDNLPIDQDY 128
 DB 77 QAMRLVEEINNSTALNITLGLYDVCSDSANVYATRLVSLPGOH--IELQDDL 133
 QY 129 SNYISRVAVVIGPDNSESMVYANFLSLFLPOLYTSASIDELRKVPALRLTPPSAD 188
 DB 134 LHTSPVLAIVGPDSTNNAATTAALLSPFLVPMISTAASETLSVKROPYSLRTIPMDK 193
 QY 189 HHVEAMVOLMIFRMWMTIVVYSDPYGRDNGOLLGEEVARRDICIARQETLP-TLQPNQ 247
 DB 194 YQVETMVLLOKFGMTWISLVGSDQGLQVLALENQATGGCICIAKIDIMFPAQVGD 253
 QY 248 NMTSEERRLTVYDKLOQSTARVYVYSPDLTYHFNEVETLRQFTGAVMYASSAID 307
 DB 254 -----ERMOCMLRHLAQAGATVYVVSQRLARVFESVVLNLTGKVAWVASAMALS 306
 QY 308 PVLHNLTELGHTGFTGTTIOSVPIPGFSSEFEMQPA---GPPLSRTSOSYTCNOBCD 364
 DB 307 RHITGVPSIQIRIGWYLGVAIGRAVPLGAKPEBARARADKREPRCHGSMC-SSNOJCR 365
 QY 365 NCLNATLSEFNTILRLSGERVYYSVYSAVYAAHALHSLGCDKSTCTRRVYYPWQLEBI 424
 DB 366 EQQAFMATMPRLKAFSSMSAVNAVYAVAAHGLHQLGCSAGCSGRVYPMQILROI 425
 QY 425 WKVNFLLDHOIFPDPODVALHLEIVOMQDRSQNPQVAS--YYLQRLQKNIQDIS 482
 DB 426 HKVHFLKRDVYAFENDNDPLSSYNIAMDWNGKPMFTFVLGSSYWSVQOLNI-NETRIQ 484
 QY 483 HTVANNITPMSCSKRCOSGOKKRPVGIHVCCFECIDCLPGFELNHTDEDECOACPNNE 542
 DB 485 WHGKNOVPKSYCSDCLBEGHQRVVTGHHCCFECVPCGAGFLNKS-DLNRQPCGKEE 543
 QY 543 WSYOSETCFKROLVLEHMEAPITAVALLAAGFLSTLALIVFWRHQPTIVSAGP 602
 DB 544 WAPESGQCFPRVTVFALREHTSMVLLAANTLLLLLLGTLGFLAMHLDPPVRSAGR 603
 QY 603 MCFMLTLILVAVVVPVYVGPVKSTCLCQALPPLCFTTICISIAVRSFOIYCAFMA 662
 DB 604 ICFMLGSLAAGSGSLVGFEGFPTRACLLRQALFALFTFLSCILTVRSFOLLITIFRS 663
 QY 663 SRFPRAYSYWRVYGPVYSMAFIVLAKVIYIGMLARPOSHPRTPDDPKITIVSCPN 722
 DB 664 TKVPTFYHAWQNHGAGLEFVMISSAQLLCLTLVWVTPPAREYORFPLVMECTET 723
 QY 723 YRNSL-----LFTNSLDLLSVGFSFAYMGKELPTYNNAKFTTSMTEYFSSVSLC 776

AAE11970
ID AAE11970 standard; Protein: 763 AA.
XX
AC AAE11970;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human novel G-protein coupled receptor (NGPCR) protein #2.
XX
KW Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
KM antisense-therapy; signal transduction; behavioural disorder; obesity;
KW heartbeats; inflammation; immune disorder; diabetes; cancer;
KM coronary disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT MISC-difference 29 /note= "Encoded by TYT"
FT MISC-difference 294 /note= "Encoded by RCA"
FT
PN MO200172842-A2.
PD 04-OCT-2001.
PE 28-MAR-2001; 2001WO-US09996.
PR 28-MAR-2000; 2000US-192978P.
PA (LEXI-) LEXICON GENETICS INC.
PI Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abulin A;
PI Zambowicz B, Sands AR;
XX WPI: 2001-616474/71.
DR N-PSDB: AAD19502.
XX
PT Novel isolated polynucleotides encoding human G protein coupled
PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
PT expressed NGPCRs for diagnosis of disease, and as probes or primers -
XX
PS Claim 4; Page 75-76; 80pp; English.
XX
CC The present sequence is human novel G-protein coupled receptor
CC (NGPCR) protein. NGPCRs are transmembrane proteins that span
CC the cellular membrane and are involved in signal transduction after
CC ligand binding. The NGPCR polynucleotide sequences, are useful
CC in diagnosis and treatment of a disease involving NGPCR, for detecting
CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NGPCR in
CC the body or abnormalities in the signal transduction pathway mediated by
CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
CC trial monitoring and/or the treatment of physiological (heartbeats rate)
CC or behavioural disorders. NGPCR is useful for identifying compounds
CC useful in the therapeutic treatment of obesity, inflammation, immune
CC disorders, diabetes, heart and coronary disease, metabolic disorders and
CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
CC genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests.
XX
SO Sequence 763 AA;

Query Match 28.4%; Score 1260.5; DB 22; Length 763;
Best Local Similarity 37.0%; Pred. No. 3.1e-112;
Matches 288; Conservative 124; Mismatches 331; Indels 35; Gaps 12;

QY 75 MRFAVEEINNNSSILPGVILGELYVDYCYISNNQPVLYFLA----HEDNLIPIQEDYSN 130
DB 1 MRGVEEINNNSSILPGVILGELYVDYCYISNNQPVLYFLA----HEDNLIPIQEDYSN 130

QY 131 YISRVAVIGPDNSSESVTFANPLSLPOTITSATISDELDRKVRPALRTSPADNH 190
DB 58 YSPYLAIVIGPDNSSESVTFANPLSLPOTITSATISDELDRKVRPALRTSPADNH 117
QY 191 VEAMVQMLHFRMWMVIVLVSSDFYGRDNGQLGERVARRDICIAFOETLP-TLQPNOM 249
DB 118 VETWVLLQKFGWVWISLVSSDDYGGQVQALENGATGOGICIAFKIMFMSQVGD- 175
QY 250 TSEERQRLVTVYDKIQOSTANVAVVFPDLTLHFNEVLENONFTGAVIASESMAIDPV 309
DB 176 -----ERMOCIMRLHAQAGAVVVVFSRQARVFEESVVLNTMGKVVASEAMALSRH 230
QY 310 LHNLTGHLGTEPFGITQSPVIGPGESEFRMGPA--GPPLSRSTQSTQNOEDNC 366
DB 231 ITGVPGLQRIQMGVGLVALQKRAVPELKAEEYAVARADKEARPCPKSGWC--SSNQLREC 289
QY 367 LNAITSEFTILRLSGERVVSVYSAVYVAHALSHLSCDCKSTCKRVVYPMOLEETWK 426
DB 290 QAFNAHMPKLFKAFMSMSANAVYAVVAGHQLGLGACAGACSRGRVYPMOLEETWK 349
QY 427 VNFLLDHIQIFPDQGVVLAHLEIVQOMDRSONPQSVAS--YPLQROLKNTQDISWH 484
DB 350 VHFLLHKDVAFNDNRDPLSSYNIIAMDNGPKTFVILGSSFTWSPVQVLT-NETKIQWH 408
QY 485 TVNNTIIPMSCKSKQSGQKKKPVGIHVCEPCIDCLPGFTLNHTEDEYEGOACPNNEWS 544
DB 409 GKNOVPRKSVCSDDLSEHQRVVTGFHHCCEYPCGAGTFLNLS-DLYRQPCKEEPA 467
QY 545 YQSETSCFRQLVLEWHEAPLIAVALLAAGFTLTAIIIVFHRFOTPIVRSAGCPMC 604
DB 468 PEGSOTCFPRIVYFLALREHTSWVLLANTLLILLTCTAGLFAMHLDTPVRSAGRLC 527
QY 605 FLMTLLIVAVVVPVYVPPRVSTCLQALFPLCTPITCISLANVSFOYCAFKMSR 664
DB 528 FLMLSLAAGSGLYGFEPFACILRQALFALGFIFPLCSLVRSFOIILFFKSTK 587
QY 665 FPRVSYVWRVQGPVYSAFTVLMVIVIGMLARPQSHRTPDDPKRTIVSCNPYR 724
DB 588 VPTFHAMVQNHGAGLFPMISSAQLCLTWVYVWTLPRAREORFHLVMEETEF-- 645
QY 725 NSL-----LENTSIDLLSVYGSFAYMGEKELPTNYNEAKFTILSMFTYSVSLCTF 778
DB 646 NSLGFILFLYNG-----LLSIAFACSYLGRDLENYEACVFEISLIFNEVSWIAFEPT 701
QY 779 MSATSGVAVTVIDLIVYVLLNLALSLGFGPKCVMLLEPPERNPAYNSMIQGTMR 836
DB 702 ASYVDKTYLPANMMAGLSLSSGFGYFLPKCYVILICRPDLNSTEHPQASIQDYTR 759

RESULT 15
ID AAU00508 standard; Protein: 1059 AA.
XX
AC AAU00508;
XX
DT 29-AUG-2001 (first entry)
XX
DE Chicken calcium-sensitive receptor (CaR) protein.
XX
KW Avian; chicken; calcium-sensing receptor; CaR; clone CID;
KW extracellular calcium homeostasis; parathyroid hormone; PTH;
KW serum calcium regulator; bone deposition.
XX
OS Gallus sp.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT Peptide /label= Signal_peptide
FT Domain 1..611
FT Domain /label= Extracellular_domain
FT Protein /note= "Amino-terminal predominantly hydrophilic domain"
FT Protein /label= Mature_CaR_protein

```

FT Region 136.165
FT /note= "Hydrophobic region characteristic of
FT calcium-sensing receptors and metabotropic
FT glutamate receptors"
FT Domain 612..861
FT /note= "Hydrophobic core comprising helical
FT transmembrane domains"
FT Domain 862..1059
FT /note= "Carboxy-terminal hydrophilic domain"
XX
XX U56210964-B1.
XX
XX 03-APR-2001.
XX
XX 14-AUG-1998; 98US-0134513.
XX
XX 18-AUG-1997; 97US-0058095.
XX
XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
XX Brown EM, Diaz R, Bai M, Quinn SJ;
XX
XX WPI; 2001-289636/30.
XX
XX N-PSDB; AAS01709.
XX
XX New avian calcium-sensing receptor polynucleotide and encoded receptor
XX protein, useful for regulating serum concentration of calcium animals,
XX particularly in chickens
XX
XX Claim 1; Fig 2A-2D; 43pp; English.
XX
XX The present sequence representing an avian (chicken) calcium-sensing
XX receptor (Car) is isolated from chicken parathyroid gland cDNA clone C1D.
XX Car is involved in regulating extracellular calcium homeostasis by
XX controlling PTH (parathyroid hormone) secretion. The polynucleotide
XX encoding Car is useful for producing calcium-sensing receptor protein,
XX which can be used to regulate extracellular calcium homeostasis and to
XX regulate serum calcium levels in chickens and related species. By
XX increasing serum calcium, more rapid growth is obtained due to an
XX increased rate of bone deposition, and eggs of higher quality are
XX produced. A DNA construct comprising the Car polynucleotide is useful
XX for developing transgenic animals expressing a mutated form of the
XX calcium-sensing receptor. The Car polypeptide can be used to produce
XX antibodies to Car, which can be used to detect the presence of Car
XX protein using immunoassays. Also described are methods and compositions
XX which can be used to modulate the serum concentration of calcium in
XX humans and animals.
XX
XX Sequence 1059 AA:
XX
XX Query Match 24.9%; Score 1107; DB 22; Length 1059;
XX Best Local Similarity 31.9%; Pred. No. 3.4e-97;
XX Matches 284; Conservative 148; Mismatches 354; Indels 104; Gaps 24;
XX
XX 9 CSLFFLL--WTLAPRAESDYLPDGYLLGLFSLHAKKGIYHLNFIQVPMCKEYEVK 66
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 6 CCLILLFTWTAAYGPNQAKKKDILLGLFPIHF--GVAAKD--QDKKSRESEYEC 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 67 IGYNL-----NQARFAVEEINNDSSLPGVLLGLEYIDVC--YISNNYQVLYFLA--HED 119
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 61 IRYNFRGRRLQAMIFAEIEENNSPILLPNTLGRIPDTCTNYSKALEATLSFYAOKNI 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 120 NLLPQIE--DYSNYISRVAVIGPNSESVTVANFLSLFLLPQITYSAISDELKDYRF 177
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 121 DSLNIDECNCSHPISTIAYVAGTSGSVTAVANILGLFYIPQVSAYSSRLSNKQF 180
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 178 PALRTTPSADHYEAMQQLHFRWMIYIVSSDTYGRDNGQLGERVARDCIAFQ 237
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 181 KSFRLTIPNDHQATAMADIIIEFRMNVGTIADDDYGRGPIKFRFEAEERDIDIFS 240
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 238 ETLPTLPQNMNTSEERLRTIYDQLQOSTARVYVSPDLLYHFNENELRONFGAV 297
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 241 ELI-----SOTSDEEIQVVEVI--ONSTARVIYVSSGPDLEPLIKETIRKNTIGKI 292

```

```

QY 298 WIAESNAIDPVLNLTGLTFLGTTIOSVPIPGFSERRENGPOAGPPLSRTSOSY 357
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 293 WIAESNAISSSLIMPEFFRIVISTIGPALKAGIIPGRRELIQ--KVHPKKSANNQFAK 349
QY 358 TCNOECNCL-----NATLSFN----- 374
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 350 EFWEETFCYLPSESKNSPASASFHKAHEBGLGAGNGTAAPRPPCTDENTISYETPYMD 409
QY 375 -TILRLGERVYVSAYVAVAHALSLIG-----DKSCTKRVYPMQOLLEIM 425
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 410 FTHLRIS-----YVYLAIVSIAHALODITYCPGKGLFTNGSCADTKKQAMQVLNHR 464
QY 426 KVNFTL-LDHOIFEDPDQVALHLEIYOWOMDRSONP--FOSVASYPL-----OROLNI 478
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 465 HLNFTSMNGEVDDEGDIVGNYSIINMHLSPEDGSVFEVGHYVYAKKGERLINE 524
QY 479 QDISMHTYVNTIPMSKSKRQSQOKKPV-GIHYCCPECIDCLPGFPLNTEDEYEQQA 537
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 525 NKILMSGFSKEVPSPNSRDLPGTRKGLIEGEPTCCFECVDCPDGEYSDET--DASACDK 583
QY 538 CPNNEMSYOSETSCFRQOLVLEWHEAPTIYAVALLAALGFLSTLAIIVIEFRHPQPIVR 597
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 584 CPEDYMSNENTSCIPQIEFLSWTERPGIALTLFAVLGIFLTSFVLGVTFKRNPTIVK 643
QY 598 SAGPMPCLMTLLLVAVMVVYVGPVSTCLCQALPFLCFTTICISCIABRSQIVC 657
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 644 ATNRELSTYLLFSLICFSSSLFEGEPQNTCRLOPAGISFVLICISCLIVKTNHVL 703
QY 658 AFKMASRPRPA-YSVWRYQGPVSMATVYLVKAVIYVIGMLAPQSHRPDPDPKITI 716
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 704 VFE--AKPTSLHKKMWGLNQLFLVLCITFVOIVICVIMLYTAPSSYRNHELEDTIF 761
QY 717 VSCNPNTNSILFNTSLDILLVGVGFAYWKGELPTNYNEAKFTLSMTEFTSYSLC 776
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 762 ITCHBGSLMAGFLIGTCLLAALCFPAFASKRLPENEAKTITTSMLFTIYWS-- 819
QY 777 TFMASVGLVTIVDLVTYVNLALASLG---YFGPKCYMILPEPNT 822
QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 820 -FIPAVASTYGVFS-AVEYAIILAASBGLLACIFPNKVTYILKPSRNT 867

```

Search completed: May 19, 2003, 09:49:59
Job time : 41.6935 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:47:40 ; Search time 19.8343 Seconds
(without alignments)
4061.679 Million cell updates/sec

Title: US-09-927-315-9

Perfect score: 4443

Sequence: 1 MGPRKATGSHFLMWLAEE.....ERNTPAYENSMIGYTMRD 838

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR_73:*
2: pir1:*
3: pir2:*
4: pir3:*
5: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1094	24.6	1078	2	AS6715
2	1088.5	24.5	1079	2	I59362
3	1087.5	24.5	1085	2	SA0476
4	1079	24.3	1088	2	B56715
5	1012	22.8	858	2	JC7683
6	724.5	16.3	915	2	A49874
7	721.5	16.2	912	2	JH0563
8	718	16.2	879	2	JH0562
9	718	16.2	879	2	JC7160
10	699.5	15.7	872	2	JH0561
11	685	15.4	1171	2	A42916
12	675.5	15.2	1180	2	JC2132
13	675.5	15.2	1212	2	JC2131
14	673	15.1	871	2	A46742
15	672	15.1	908	2	I49142
16	655.5	14.8	999	2	I77628
17	654	14.7	1199	2	A41939
18	649	14.6	1218	2	G71376
19	632.5	14.2	1267	2	T21340
20	423	9.5	551	2	T30806
21	218.5	4.9	921	2	T51136
22	218.5	4.9	923	2	E84732
23	215	4.8	1039	2	T45779
24	200.5	4.5	953	2	E84732
25	200.5	4.5	1099	2	T16283
26	186	4.2	950	2	T51134
27	185.5	4.2	925	2	T06128
28	185	4.1	975	2	AA4550
29	184	4.1	925	2	T51133

30	184	4.1	951	2	T51132
31	177.5	4.0	1005	2	S33525
32	176.5	4.0	960	2	JF0356
33	176	4.0	933	2	C96495
34	174.5	3.9	912	2	T51131
35	169.5	3.8	941	2	T51135
36	168.5	3.8	962	2	D86186
37	160.5	3.6	1679	2	T15968
38	158.5	3.6	976	2	T51137
39	158	3.6	997	2	S33754
40	144	3.2	918	2	I58178
41	144	3.2	949	2	S19808
42	143	3.2	965	2	I51244
43	131.5	3.0	1125	1	ORURCP
44	130.5	2.9	938	2	AA6612
45	129.5	2.9	885	2	JN0335

ALIGNMENTS

RESULT 1

AS6715
calcium receptor (clone pPCAR-4.0) - human

C:Species: Homo sapiens (man)

C:Date: 19-Oct-1995 #sequence revision 19-Oct-1995 #text change 01-Dec-2000

C:Accession: AS6715; S49341; A49419; B49419; C49419

R:Garrett, J.E.; Capuano, I.V.; Hamerlind, L.G.; Hung, B.C.P.; Brown, E.M.; Hebert, J. Biol. Chem. 270, 12919-12925, 1995

A:Title: Molecular cloning and functional expression of human parathyroid calcium rec

A:Reference number: AS6715; MUID:95279439; PMID:7759551

A:Accession: AS6715

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1078 <GAR>

A:Cross-references: GB:U0759; MID:9683744; PIDN:AAA86503.1; PID:9683745

R:Pearce, S.H.S.; Thakker, R.V.

A:Reference number: AS6715; MUID:95279439; PMID:7759551

A:Accession: AS6715

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-180, 'Q', 182-989, 'R', 991-1078 <DEA>

A:Cross-references: EMBL:X81086

R:Pollak, M.R.; Brown, E.M.; Chou, Y.H.; Hebert, S.C.; Marx, S.J.; Stelmann, B.; Lev

Cell 75, 1297-1303, 1993

A:Title: Mutations in the human Ca(2+)-sensing receptor gene cause familial hypocalci

A:Reference number: A49419; MUID:94094324; PMID:7916660

A:Accession: A49419

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 178-192 <POL>

A:Experimental source: family N

A:Note: sequence inconsistent with nucleotide translation

A:Note: sequence modified after extraction from NCBI backbone

A:Note: 186-Arg mutation is associated with familial hypocalcemic hypercalcemia and

A:Note: sequence extracted from NCBI backbone (NCBIN:142453)

A:Accession: B49419

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 289-303 <PO2>

A:Experimental source: family E

A:Note: sequence modified after extraction from NCBI backbone

A:Note: 298-Lys mutation is associated with familial hypocalcemic hypercalcemia and

A:Note: sequence extracted from NCBI backbone (NCBIN:142453)

A:Accession: C49419

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 788-802 <PO3>

A:Experimental source: family J

A:Note: sequence modified after extraction from NCBI backbone

A:Note: 796-Tyr mutation is associated with familial hypocalcemic hypercalcemia and

A:Note: sequence extracted from NCBI backbone (NCBIN:142457)

C:Keywords: glycoprotein; receptor; transmembrane protein

Query Match 24.6%; Score 1094; DB 2; Length 1078;

Best Local Similarity 31.5%; Pred. No. 1.6e-75; Mismatches 342; Indels 106; Gaps 26;

Matches 279; Conservative 160; Mismatches 342; Indels 106; Gaps 26;

QY 16 WIAEPAENSDYLP-----GDYLLGGFLSHANKGIYHNLQVPMCKEYKVG 68
 Db 8 WLLATLWHTSAYGPDORAKKGDIIIGLGPPIHF---GVAAKD--QDLKSRPSEVICR 62

QY 69 YNL-----MQMRAVEEINNDSSILPGVLLGLEYIVYC-YISNNQPVLYFLA-HEDNL 121
 Db 63 YNFERFWLQMTAFIEINSSPALLPMTLGYRIPTCNTVSALERTSFEVQNNIDS 122

QY 122 LPIDF--DYSNYISRVAVIGPDNSESVMTANFLSLFLPQIYSAISDELARKVERPA 179
 Db 123 LNLDFPCNSEHISTIAVVGATGSGVSTANANLGLFYIQVYASSRLSKNQPKS 182

QY 180 LKRTTSPADHVEAMVOLMHEFRNNIIVLVSSPTYGRDNCQLIGERYARDICTAFOET 239
 Db 183 FLRTIPNDEHQATAMADIEYFRNNWVGTIADDDYGPRIEKREREAEEDICIDSESL 242

QY 240 LPTLQPNQMTSEERORLYTVDKLOOSTAVVYVFPDILLYHFNVLKQNTGAVMI 299
 Db 243 I-----SOYSDEEIQHVEYI---QNSTAKVIYVSSGPDLEPLIKETIRNITGKIWL 294

QY 300 ASESMAIDPVHLNLTGLHGTFLGITIQSVPIGPFSEF-----REMGPO 344
 Db 295 ASEMASSSLIAMPQYFHVVGITGFPALAKQIIGFREFLEKVKHPRKSVHNGAFKEFME 354

QY 345 A-----GPPPLSTQSITYCNOECNCLNATISFNTILSGE-----382
 Db 355 TFNCHLOEGANGPLVPVTFLEIGH--ESGDRFSNNSSTAFRL--CTGEMNISVETPYID 410

QY 383 -----RVVYSAYAVAAHALSHLGC-----DKSTCKRIVYPMQLEETMKNYT 430
 Db 411 YTHIRISYNYLAVYSTAHALQDIYTCLPGNGLETNSCADIKVEAMQVLAHRLHNFY 470

QY 431 -LDDHQIFPDQDVALHLEIVQMOMDRSONP--FQSVASYPL-----OROLKNIDIS 482
 Db 471 NNMGEQVTFDECGDLVGNYSIINWHLSPEDGSIVKEY-GYVNYAKKGERLFINKEKIL 529

QY 483 WHYNNNTIPMSCKRQSGQKKPV-GIHYCCFECIDCLPGLTNTEDETECOACPN 541
 Db 530 WSGSREVPSCSRDLGAGRKGIIEGEPCECEVCPEGEYSDETF-DASACNKCPDD 588

QY 542 EMSYQSECFKROLVLEHMEAPLIAVALLAALGPLSTAILIYFMHQPPTIVRSAG 601
 Db 589 FMSENHTSCIAKEIEHLSMTEPEFGIALITLFAVIGITLAVLVLFKRNTPYAKATNR 648

QY 602 PMCFMLTLVLVAVVVPVVPKPVSTCQALFPLCFCTICISCIARVSFOIVCAF-- 659
 Db 649 ELSTLLFLSLCFSSSLFIFGERQDMTCRLROPAGISFLCISCIIVKNNRLVAVEA 708

QY 660 KMASRFRPAVSYWRVQGPYSMAFITVLAQVIVIGMLARPQSHPTDPDKITTVSC 719
 Db 709 KIPISFHRK--MMGLNQLFLVLCFQMOLIVCIMLYTPAPSSYRQOELEDEIITTC 765

QY 720 NPNRNSILFNTSLDILLVYVGFPAWKGELPTNYNKAKITLTMFTFYSLSCTEM 779
 Db 766 HEGSIALMGLIGITCLLAACFFAFKRSKRLPENENAKITISMLIFETVWIS---FI 822

QY 780 SAYSGVLTVIIDLTVVNLMLAISLG-----YFGPKCYMILEYPERNT 822
 Db 823 PAYASTGKFEYS-AVEVIALIILASGELLACIFFENKIYITILEKPSRNT 868

RESULT 2

calcium/polyvalent cation-sensing receptor precursor - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
 C/Accession: I59362; A55594

R;Ruat, M.; Molliver, M.E.; Snowman, A.M.; Snyder, S.H.
 Proc. Natl. Acad. Sci. U.S.A. 92, 3161-3165, 1995
 A>Title: Calcium sensing receptor: molecular cloning in rat and localization to nerve
 A;Reference number: I59362; PMID:7724534
 A;Accession: I59362
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-1079 <RES>
 A;Cross-references: EMBL:U02889; NID:g790578; PIDN:AA052195.1; PID:g790579
 A;Keywords: calcium; glycoprotein; phosphoprotein; transmembrane protein
 F;1-20/Domain: signal sequence #status predicted <Sig>
 F;187-212/Region: hydrophobic
 F;613-635/Domain: transmembrane #status predicted <TM1>
 F;650-670/Domain: transmembrane #status predicted <TM2>
 F;683-700/Domain: transmembrane #status predicted <TM3>
 F;725-744/Domain: transmembrane #status predicted <TM4>
 F;770-790/Domain: transmembrane #status predicted <TM5>
 F;806-828/Domain: transmembrane #status predicted <TM6>
 F;841-860/Domain: transmembrane #status predicted <TM7>
 F;90,261,287,386,468,488,594,893,1005/Binding site: carbohydrate (Asn) (covalent) #st
 F;784/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status pred
 F;899,901/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status pred

Query Match 24.5%; Score 1088.5; DB 2; Length 1078;
 Best Local Similarity 31.7%; Pred. No. 4.3e-75; Mismatches 328; Indels 121; Gaps 31;
 Matches 284; Conservative 164; Mismatches 328; Indels 121; Gaps 31;

QY 10 SLFLLVTLAEPANSDYLP-----GDYLLGGFLSHANKGIYHNLQVPMCKEY 62
 Db 9 ALLALAW-----HSSAYGPDQRAKKGDIIIGLGPPIHF---GVAAKD--QDLKSRPE 56

QY 63 EYKVIYGNL-----MQMRAVEEINNDSSILPGVLLGLEYIVYC-YISNNQPVLYFLA 116
 Db 57 SVECIIRYNFGEFWLQMTAFIEINSSPALLPMTLGYRIPTCNTVSALERTSFEVA 116

QY 117 -HEDNLPIDF--DYSNYISRVAVIGPDNSESVMTANFLSLFLPQIYSAISDELARD 173
 Db 117 QNKIDSLNDFECNSEHISTIAVVGATGSGVSTANANLGLFYIQVYASSRLSKN 176

QY 174 KVRPALLRTTSPADHVEAMVOLMHEFRNNIIVLVSSPTYGRDNCQLIGERYARDIC 233
 Db 177 KNQKSLFRTIPNDEHQATAMADIEYFRNNWVGTIADDDYGRPIEKREREAEEDIC 236

QY 234 IAPQETPLQPNQMTSEERORLYTVDKLOOSTAVVYVFPDILLYHFNVLKQNTGAVMI 293
 Db 237 IDESELI-----SOYSDEEIQHVEYI---QNSTAKVIYVSSGPDLEPLIKETIRNIT 288

QY 294 TGAVWISSEMAIDPVHLNLTGLH-GTFLGITIQSVPIGPFSEF-----338
 Db 289 TGRWTLASEMASSSLI-AMPEYFHVVGITIGRLAKQIIGFREFLEKVKHPRKSVHNG 347

QY 339 -REMGPOA-----GPPPLSTQSITYCNOECNCLNATISFNTILSGE-----382
 Db 348 AKFWEETFNCHLOEGAKGPLVPVTFVRSH--EEGGRLLNSSTAFRL--CTGEMNIS 403

QY 383 -----RVVYSAYAVAAHALSHLGC-----DKSTCKRIVYPMQLEETMKNYT 423
 Db 404 VERPEYMDYEHRLRISYNYLAVYSTAHALQDIYTCLPGNGLETNSCADIKVEAMQVLA 463

QY 424 IMKRNFT-LDDHQIFPDQDVALHLEIVQMOMDRSONP--FQSVASYPL-----OROL 475
 Db 464 LRHLNFTNNMGEQVTFDECGDLVGNYSIINWHLSPEDGSIVKEY-GYVNYAKKGERL 522

QY 476 KNIDISWHNTNTPMCMSCRCOSGOKKRPV-GIHVCEECIDCLPGLNTHEDYE 534
 Db 523 INEKLIMSGSRREVPENCSDOAGTRKGIIEGEPTCCBCEVCPGEGSGER-DASA 581
 QY 535 COACPNNMWSYQSTSCFRQJLVEHWEAPLIAVALLAALGSLTALIVFWRHQTP 594
 Db 582 CDKPPDDFMNSNENHTSCIAKEIEFLIATEPFGIALTLFAVLGIFLAVLGIFKFRNT 641
 QY 595 IVRGAGGMCMLMTLLILVAIVNVVYVGGPKVSTCLCRQALPFLCFTICISCIANSFQ 654
 Db 642 IVKANKRELSTLLSLCCFSSSLFEGEPQDWCRLRQPAFGISFVLCISCIILVKNR 701
 QY 655 IVCAF--KMASRPRAYSVWRYOGPYSAFIVLKAVIYIGMLARQSHPRTPDDP 712
 Db 702 VLVFEAKIPTSPHKK---WMLNLOFLVFLCTFMQILICIMVXTAPPSYRNHELD 758
 QY 713 KITVSCPNRNLSLFTSL---DLISVGFSPAYNGKELPTNYNEAKFTTSLMTEYF 769
 Db 759 EIIETTC---HEGSLMAGSLIGYTCLLAICFFPAFKSRKLPEFNFAKFTTSMLEIF 815
 QY 770 TSSVLCFMSAGVLTIVDLVTVNLALISIG---YEGPKCYMILFPERNT 822
 Db 816 IWMIS---FIRAYASTYKFPVS-AVEVAILAASFGLLACIFPNKVYIILFKPSRNT 868

RESULT 3

S40476
 Ca(2+)-sensing receptor - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Nov-1999
 C:Accession: S40476
 R:Brown, E.M.; Gamba, G.; Riccardi, D.; Lombardi, M.; Butters, R.; Kifor, O.; Sun, A.; H
 Nature 366, 575-580, 1993
 A:Title: Cloning and characterization of an extracellular Ca(2+)-sensing receptor from h
 A:Reference number: S40476; MUID:94077182; PMID:8255296
 A:Accession: S40476
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1085 <BRO>
 A:Cross-references: GB:567307; NID:g453108; PIDN:AAB29171.1; PID:g453109

Query Match 24.5%; Score 1087.5; DB 2; Length 1085;
 Best Local Similarity 31.6%; Pred. No. 5.2e-75;
 Matches 279; Conservative 161; Mismatches 347; Indels 97; Gaps 28;

QY 11 LEFLMVALEPAENSDFLPGDYLLGGLFSLHANKGIVHNLFLQVP---MCKEEVAVI 67
 Db 11 LAFSTWCTSAIGPDORAKKGDIIIGLPIHFGV-AVKDDLSRPSVEICIRNFR-- 67
 QY 68 GYNLMQARFAVEEINNDSSLLPGVLLGLEYIVDC-YISNVQPVLYELA-HEDNLEPIQ 125
 Db 68 GFRMLQAMIFAEIENSSPALLPMTLGYRIEDTCNFSKALEATLSVAQNKIDSLMD 127
 QY 126 E-DYSNISRVAIVIGDNSESWTANFLSLFLPQITTSALSDERDKVRPALLRT 183
 Db 128 EFCCKSEHPISTIAVVGATGSGISTAVANLLGLFYIPQVSYASSSRLSNKNOFSFLRT 187
 QY 184 TPSADHVEAVQMLAHRMMWIIYLVSSDTYGRDNGQLGGERVARRDICAFOETLPTL 243
 Db 188 IPNDEHQATAMADIIIEYRMMWVGIIAADDDYGRGRIKFRFEAEERICIDFSLIT--- 244
 QY 244 QPNQMTSEERQRLVTVDKLQOSTARVYVSPDLTLYHFEENEYLRNFTGAVWIAS 303
 Db 245 ---SQYDE-EKIQOYVEVIONSTAKIVIVFSSGPDLEPIKELIVRNITGRIMLAS 299
 QY 304 WAIDPVHLNLELGH-LGTFIGITIQSPVPIGFSF-----REMGPOA-- 345
 Db 300 WASSSLI-AMPEYFHVGGTIGFGKAGQIGCFREFLOKVRPRKSVHNGAKFEWEETFN 358
 QY 346 -----GPPPL-----SRISQSYT-----CNOCNDCLNLTLSFNTILRISG 381
 Db 359 CHLOGGAGPLPVDITFLRGHEGGARLSNSPTARPLCTGE-ENISSVETIPMDYTHL-- 415

QY 382 ERVYSYSAVAVALHLSLGC-----DKSTCKRVRYPWQLEIKRVNFTL-L 432
 Db 416 -RISNVLAYIASIAHALODITYCIPKGLFNGSCADIKKVAQVLAHLHLFTSM 474
 QY 433 DHOIFPDQGDVALHLEIVQWQDRSONP--FQSAVSYPL-----QRLKNIDISWHT 485
 Db 475 GEQVTFDECGDLGAGNYSIINHLSPEDGSIVYKEV-GYNNYAKKGERLFINDEKLMSG 533
 QY 486 VANTIPMCMSCRCOSGOKKRPV-GIHVCEECIDCLPGLNTHEDYEQACQNNMWS 544
 Db 544 FSREVPYVWRYOGPYSAFIVLKAVIYIGMLARQSHPRTPDDPKITTIYSCPN 592
 QY 545 YQSESCFRQJLVEHWEAPLIAVALLAALGSLTALIVFWRHQTPYRSAGGPMC 604
 Db 593 NENHTSCIAKEIEFLIATEPFGIALTLFAVLGIFLAVLGIFKFRNTPIYKARNR 652
 QY 605 FLMTLLVAIVNVVYVGGPKVSTCLCRQALPFLCFTICISCIANSFQIYCAF--KWA 662
 Db 653 YLLFSLCCFSSSLFEGEPQDWCRLRQPAFGISFVLCISCIILVKNRVLVEAKIP 712
 QY 663 SREPRAYSVWRYOGPYSAFIVLKAVIYIGMLARQSHPRTPDDPKITTIYSCPN 722
 Db 713 TSEHKK---WMLNLOFLVFLCTFMQILICIMVXTAPPSYRNHELDLITTCB 769
 QY 723 YKNSLFTSLDLISVGFSPAYNGKELPTNYNEAKFTTSMTEYTSVLCFMSAY 782
 Db 770 SLMAIGFLIGYTCLLAICFFPAFKSRKLPEFNFAKFTTSMLEIFIWMIS---FIRAY 826
 QY 783 SGVLTIVDLVTVNLALISIG---YEGPKCYMILFPERNT 822
 Db 827 ASTYKFPVS-AVEVAILAASFGLLACIFPNKVYIILFKPSRNT 869

RESULT 4

B56715
 calcium receptor (clone phpcar-5.2) - human
 C:Species: Homo sapiens (man)
 C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 21-Jul-2000
 C:Accession: B56715
 R:Garrett, J.E.; Capano, I.V.; Hammerland, L.G.; Hung, B.C.P.; Brown, E.M.; Hebert,
 J. Biol. Chem. 270, 12919-12925, 1995
 A:Title: Molecular cloning and functional expression of human parathyroid calcium rec
 A:Reference number: B56715; MUID:95279439; PMID:7759551
 A:Accession: B56715
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1088 <GAR>
 A:Cross-references: GB:020760; NID:g683746; PIDN:AAA86504.1; PID:g683747
 C:Keywords: glycoprotein; receptor; transmembrane protein

Query Match 24.3%; Score 1079; DB 2; Length 1088;
 Best Local Similarity 31.1%; Pred. No. 2.3e-74;
 Matches 279; Conservative 160; Mismatches 342; Indels 116; Gaps 27;

QY 16 WYLAEPANSDFYLP-----GDYLLGGLFSLHANKGIVHNLFLQVPMCKEYEVYIG 68
 Db 8 WYLAALTWHTSAYGPDORAKKGDIIIGLPIHF---GVAKD--QDLKSRESVEICIR 62
 QY 69 YNL-----MQARFAVEEINNDSSLLPGVLLGLEYIVDC-YISNVQPVLYELA-HEDN 121
 Db 63 YNFRGFRMLQAMIFAEIENSSPALLPMTLGYRIEDTCNFSKALEATLSVAQNKIDS 122
 QY 122 LPIDQ--DYSNISRVAIVIGDNSESWTANFLSLFLPQITTSALSDERDKVRP 179
 Db 123 LINDFCNCSEHPISTIAVVGATGSGISTAVANLLGLFYIPQVSYASSSRLSNKNOFS 182
 QY 180 LIRTPSADHVEAVQMLAHRMMWIIYLVSSDTYGRDNGQLGGERVARRDICAFOET 239
 Db 183 FLRTIPNDEHQATAMADIIIEYRMMWVGIIAADDDYGRGRIKFRFEAEERICIDFSBL 242
 QY 240 LPTLPQNMNTSEERQRLVTVDKLQOSTARVYVSPDLTLYHFEENEYLRNFTGAVWI 299
 Db 243 I-----SQYSDDEEIQHVEVET---QNSTAKIVIVFSSGPDLEPIKELIVRNITGRIML 294

A:Title: Cloning and expression of a new member of the L-2-amino-4-phosphonobutyric acid

A:Reference number: 157954; MUID:94195260; PMID:8145723

A:Accession: 157954

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-915 <RES2>

A:Cross-references: EMBL:006832; NID:g459657; PIDN:AAA20655.1; PID:g459658

C:Genetics:

A:Gene: MGLUR7

C:Superfamily: metabotropic glutamate receptor 4

C:Keywords: neurotransmitter receptor

Query Match 16.3%; Score 724.5; DB 2; Length 915;
Best local similarity 28.0%; Pred. No. 3e-47;

Matches 257; Conservative 152; Mismatches 353; Indels 157; Gaps 41;

```

QY 9 CSLEFLWLAEPNSDFLP-----GDYLLGLFSLHANKGIIVHLNLTQVPMCKEY 62
DB 19 CVLEVLICVLAARQGEVAPHSIRIEGDVTLGLFVPAH--KGP-----SGVP-CGDI 70
QY 63 EVKVGYNLMQAMRFAVEEINDDSLPGVLLGYEIVDVCYISNNQVPLYLEAHEDNL 122
DB 71 K-RENGIHLEALYALQDINDPMLPVTLGARLIDTC--SRDTYALEQSLTFYQAL- 126
QY 123 PLOEDVSNY-----ISRVAVIGPDNSSESVMTVANPLSFLPQITYSASIDP 170
DB 127 -IQKDSIDRCINGEPPEVVEKQVGVYAGSGSSIVANILRLFOIPQISTASTAE 185
QY 171 LNDKVFAPLLRTTSDADHVEAMVQMLHFRNMIIIVLVSDTYGRDNGQLGERVAR 230
DB 186 LDDRRYEDFSRVPDPSFOAQMVDIVKALGMNVSTLASBSYG-EGVESFQISKE 244
QY 231 --DICAFPEITLPTLOPNMTSEERQV----TIVDLQOSTARVAVYFSDTLIYAF 284
DB 245 AGGLCTAQSRIIP-----QERKRTIDPRITKQLLDPPNSAAVVIYFANDEDIKOI 295
QY 285 FNEVLNQNTG-AWMIASEMA--IDPVHLNLTGLTFLGTTIO--SVPIGSEPR 339
DB 296 LAAKRAQDVGHLWVSGDSGSKINP-LHQHEDIAE---GATIQPRATVEGDADF 350
QY 340 ENGPRQAPPLRTSOSTYCN-----QECNDC--LNAITSEFTILSG-ERY---- 384
DB 351 -----TSRTLENNRNRMFAEYWEFNCKLTIGSKKEDRCKTQGERIGKDS 400
QY 385 -----VYSVSAVYAVAHLSLLGDSCTK-RVYRP-----MOLLEIKVNF 429
DB 401 NVEQEKQVGLDAYAYAAHAAHNM--NKDLCADRYGCEPEQAGGKLKYLINAVNF 457
QY 430 T-LLDHQIFPDQGVALLHLEIVQWQDRSONPFQSVASYPLQROLKNIODISMTVNN 488
DB 458 NSAGTPVPMENKNGDAPGRYDIFQYQTNTNPGYRLIGQWTDQL-NIEMQMGKVR 516
QY 489 TTPMSKCRQSGQKKKRVGIIHCCFECIDCLPGLNHTDEVEYEQACPNNEKSYGE 548
DB 517 EIPSSVCTLPCKPQGRKTKQKGPCCWCEPC--DGYOYODEMTCQCHPYDQRENER 573
QY 549 TSCFQKQVLEEMHEAPTLAVALLAAGFLSTLAILIVIRHQRPIVVSAGGPMOFLM 608
DB 574 TGCQNIPIIKLEHSHMAVYIPFLAALGITATLEVATIRINDPIVVASGELSYALL 633
QY 609 TLLAVAMVVPYVGPVKSTCLRQALPLCTTICISIAV-----RSFOIYCAFKMAS 663
DB 634 TGLPICTIITFLMIAPKPDVAVCSFRVFLGL--GMCISYALLTKNRIRYRIFEGQKRV 691
QY 664 RPPRAVSTVWRQGPVSNAFITV-LKAVIIVIGMLARQSHPRIDDPDKTTI-----V 717
DB 692 TAPRLISPTSLA--ITSLISVOLGLVIFWGV-----DPPNIIIDDEKH 736
QY 718 SCNPYRNSILFNTSLDL-LLSVGFSS-----FAVMGKELPTNVNEAKFITLSMTF 767
DB 737 TANPEQARVYKLCDDITDLDIGSLGSIILAMVTCYIAKTKGVPENNEAPRIGFTM-- 794
QY 768 YFTSSVSLCTFMSAVSGVLTIVDL-----LTVLNLILA-ISLG-YFQPKCYMILYPE 819

```

DB 795 -YTCIWMIAFPPIFGTAQSAEKLYITQTTLTISMNLSASVALGMLMPKVIITIIIFPE 853

QY 820 RMT---PAYFNSMIOGYTM 835

DB 854 LNVQKRKRKRSFKAVVYATVM 872

RESULT 7

JH0563 metabotropic glutamate receptor 4 precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Jun-1992 #sequence #revision 30-Jun-1992 #text_change 17-Nov-2000

C:Accession: JH0563; 158149

R:Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.

Neuron 11, 169-179, 1992

A:Title: A family of metabotropic glutamate receptors.

A:Reference number: JH0563; MUID:92110002; PMID:1309649

A:Accession: JH0563

A:Molecule type: mRNA

A:Residues: 1-912 <TAN>

A:Experimental source: brain

R:O'Hara, P.J.; Sheppard, P.O.; Thøgersen, H.; Venezia, D.; Halldeman, B.A.; McGrane, Neuron 11, 41-52, 1993

A:Title: The ligand-binding domain in metabotropic glutamate receptors is related to A:Reference number: 158149; MUID:93332899; PMID:8338667

A:Accession: 158149

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-123, 'R', 125-912 <RES>

A:Cross-references: GB:M90518; NID:g205400; PIDN:AAA93190.1; PID:g205401

C:Comment: This protein is coupled to a G protein and evokes a variety of functions b

C:Genetics:

A:Gene: GLUR4

C:Superfamily: metabotropic glutamate receptor 4

C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; phos

F:1-33/Domain: signal sequence #status predicted <SIG>

F:33-912/Product: metabotropic glutamate receptor 4 #status predicted <MET>

F:588-610/Domain: transmembrane #status predicted <TRI>

F:625-645/Domain: transmembrane #status predicted <TRI>

F:657-675/Domain: transmembrane #status predicted <TRI>

F:700-720/Domain: transmembrane #status predicted <TRI>

F:751-772/Domain: transmembrane #status predicted <TRI>

F:786-807/Domain: transmembrane #status predicted <TRI>

F:822-844/Domain: transmembrane #status predicted <TRI>

F:981,301,454,484,569/Binding site: carboxylate (Asn) (covalent) #status predicted

F:621,689,695,859,870/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 16.2%; Score 721.5; DB 2; Length 912;
Best local similarity 26.5%; Pred. No. 5e-47;

Matches 240; Conservative 161; Mismatches 349; Indels 155; Gaps 37;

```

QY 8 ICSLEFL--WV---LAEPAN---SDFTLPDYLGLFSLHANKGIIVHLNLTQVPMCKEY 59
DB 16 ICLLLSLVAPWVPSISGKRGKHPHNSIRIDITLGLFVPHGKS-----EGKAC 67
QY 60 KEVEKVGYNLMQAMRFAVEEINDDSLPGVLLGYEIVDVCYISNNQVPLYLEAHED 119
DB 68 GEIK-KEKGIHLEALFALDRINDPILPITTLGARLIDTC--SRDTHALEQSLTFYQ 124
QY 120 NLLPQEDVSNY-----ISRVAVIGPDNSSESVMTVANPLSFLPQITYSASIDP 167
DB 125 AL--TEKDSIEVRCSGGPIITTKPERVGVYAGSGSSIVANILRLFOIPQISTASTAE 182
QY 168 SDELNDKVFAPLLRTTSDADHVEAMVQMLHFRNMIIIVLVSDTYG----- 216
DB 183 APDLSDNSRYEDFSRVPDPSFOAQMVDIVKALGMNVSTLASBSYGESEVFAFIQKS 242
QY 217 RONGQLGSRVARRODLCIAFOETLPTLOPNMTSEERQVLYTDKIQOSTARVAVYVS 276
DB 243 RENG-----GVCIAQSVKIRP-EKPTGEFPIIKRL-----LETNARKGIIITPA 285
QY 277 PDLTYHFFNEVLQNFTG-AWMIASEMAID--PVLHNLTELGHLGTFGLGITI--QSV 331

```

Db 286 NEDDIRVLEAARANOQHFFFMGSDSGSKAPVL-RLEEVAE---GAVTILPKRMS 340
 QY 332 IPGSEFEMQPOAGPPLISRT-----SOSYCNOCDCNLANTLSTFNILRSG 381
 Db 341 VRGDRKFS-----SITLNNRRNRWFAEFMEDNEFKLSRHLKSGSHKCTN 390
 QY 382 -ERV-----VYSVSAVAVAAHLSLGLCDKSTCKRY-----YYPMLL 421
 Db 391 RERIGODSAYEQEGKVGQVIDAVYAMGHALHAM---HRDLPGRVGLCPRPDPVQGLL 447
 QY 422 EIRKKNFT-LLDQIFPDQDVALHLELYQOMDSQNPQSVASTYP-LQKOLANTQ 479
 Db 448 KYIRNVESGIAGNPVTFENEGDAPGRYDIYOQLRNGSAPRYKYGSTWDLHLRIEMQ 507
 QY 480 DISHNTVNTIPMSCKRCSGOKKRVGIHCCFECIDCLPGFNLHTEDEECQAC 539
 Db 508 ---WPGSQQLPRSLCSLPCQPERKTKYKMACCMHCPC---TGIOYQVDRYTKTCP 561
 QY 540 NNEWSYSETSCFRQLVLEHMAPTIAVALLAFLSLTALVTFWRHPQPIVISA 599
 Db 562 YDMPTERTSCQPIPIYKLEMDSPMAVLPFLAVGIAATLFFVFEVRYNDPIYAS 621
 QY 600 GGPNCFLMLTLVAVMYVYVGPVKYRCLQALEPFLCTICISCIANRSQIYCAF 659
 Db 622 GRELSYLAGIFLCYATTFMLIAEPDLGTCSLRIFLGISMISYALLTKTKNIRIF 681
 QY 660 KMASRFPRAVSYYWRYOGP--YVSMAFITV-LKNAVIVIGMLARPOSHPTDPPDKITI 716
 Db 682 EQGRKRSVA-----PRFISPAQSLAITFLLISLQIGICVMWVDP-SHSVVDPPDQR--- 733
 QY 717 VSCPNRNSLLENTSIDL-LLSVYGF-----FAYNGKELPNTYNAKFTTYSMT 766
 Db 734 -TIDPRARAGLAKCDISLILCLGLSMLMTCTVYAITRGVPELFENAKIGTFM 791
 QY 767 FYFTSSVSLCTFMSAASG-----VLTVYDLVTVLNLAIISLG-YFGKCYMLIEFP 818
 Db 792 --YTTCIWMALFIPFRTSOSADKLIYQTTLTVSVLSASVSLGLMIMKXYITLIEHP 849
 QY 819 ERNRP 823
 Db 850 EQNVP 854
 RESULT 8
 JH0562
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 13-Sep-1998
 C:Accession: JH0562
 R:tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
 Neuron 8, 169-179, 1992
 A:Title: A family of metabotropic glutamate receptors.
 A:Reference number: JH0561; MUID:92110002; PMID:1309649
 A:Accession: JH0562
 A:Molecule type: mRNA
 A:Residues: 1-879 <TRAN>
 A:Experimental source: brain
 C:Comment: This protein is coupled to a G protein and evokes a variety of functions by m
 C:Superfamily: metabotropic glutamate receptor 4
 C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot
 F:1-22/Domain: signal sequence #status predicted <sig>
 F:23-879/Product: metabotropic glutamate receptor 3 #status predicted <MET>
 F:537-599/Domain: transmembrane #status predicted <TRI>
 F:614-634/Domain: transmembrane #status predicted <TRI>
 F:646-664/Domain: transmembrane #status predicted <TRI>
 F:689-709/Domain: transmembrane #status predicted <TRI>
 F:735-756/Domain: transmembrane #status predicted <TRI>
 F:770-791/Domain: transmembrane #status predicted <TRI>
 F:804-828/Domain: transmembrane #status predicted <TRI>
 F:209-292,414,439/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:610,845/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 16.2%; Score 718; DB 2; Length 879;
 Best Local Similarity 26.1%; Pred. No. 8.9e-47;
 Matches 225; Conservative 147; Mismatches 357; Indels 134; Gaps 29;
 QY 29 LPQYLLGLGFLSHANKGIYHLNLFQVPMCKEYEVVIGYNLMQMRFAVEELINDSSL 88
 Db 35 IEGDLVGLGFLPIENKGTGEEGRINEDR-----GIQRLKALFAIDELNKDNYL 85
 QY 89 LPGLVLEYEIVDVC-----YISNVQPLVFLAEHED---LLP-----IQEDYSNTISRY 135
 Db 86 LPQVGLVHILDTCSRITYALEQSLFVRASLTYVDAAEIMCPGSAIDE---NIPILI 142
 QY 136 VAVIGPDNSSVMTVANFLSLFLPQITTSYASISDELKRYKFPALLFTPSADHVEAMV 195
 Db 143 AGVIGSGSVSYSIQVANILRLFPQISYASTSAKLSDKRYDYFARTYPPDFQAKAMA 202
 QY 196 QLMHFFMNNTIYLVSSDTYGRDQGLGERVARQDICIAPQELPLIQNONMTSPERQ 255
 Db 203 ELNFFMWYTVSVASBGDYGTEGIEAFEQEARLRNICIATKRVGSRNIRKSYDS--- 258
 QY 256 RLVTYVDKLOOSTARVVVF---SPDLTLHFENEVLRONFTGAVWJASMAIDPVLHN 312
 Db 259 ---VIRELLQKPNARVVVLFMRSDSRELIJAANRV-NASFT---WVASQMGAGRESIVK 311
 QY 313 LTELGHILGTLGITIOSVPI-----PGSEFREMGPQAGPPLISRTS 354
 Db 312 GSEHVAAGA-ITTELASHPWPQFDRYPSLNPYNNHNPFRFWE----- 356
 QY 355 QSYTCN-----QEDCNCLNLTSPNTILRLSGEVRVYVSVYSAVAVAAHLSLGLG-- 404
 Db 357 QKFCSLQNKRNHRQVCDKHLAIDS-----NYQESKIMEVYNAVYAAHHLKMQRTL 411
 QY 405 CDKST--CTKRVVYPMQL--EEIKVNFVL-----LDHQIFPDQDVALHLELYQW 453
 Db 412 CPNTTKLDAMKILIDGKLYEYLLKINFAPENPKAGDSIYKFDFTGGMGKRVNENL 471
 QY 454 QMRSQNPQSVASYIYLOQLKNIDISMTHTVNTIPMSCKRCSGOKKRVGIHVC 513
 Db 472 QQTGGKYSYLKVGHW--AEFLSLDVDSIHMS--RNSVPTSQSDPCAPNEMKKNQPDVC 527
 QY 514 CFECIDCLPGFNLHTEDEECQACPNNEWSYSETSCFRQLVLEHMAPTIAVALLA 573
 Db 528 CWITICPEPEYEL---VDEFTCMDCGPGQWPTADLSQYILNPEYIMEDMAIIGPTTIA 584
 QY 574 ALGFLSTLAILVFMWHFQPIVRSAGPMPCLMLTLVLAIVVYVYVGPVKYSTCLR 633
 Db 585 CLGFLCICYITVYIKRNNPPLVYASGRELCYILLFVSLSYCTMFFIAKPSVICALR 644
 QY 634 QALPFLCTICISCIANRSQIYCAF---KMASRFPRAVSYYWRYOGP---YVSMAFIT 686
 Db 645 RLGLGTSFAICYSALIKTKCIARIFDGVNNGAQRPFIS-----PSSQVFLCLGLL 697
 QY 687 VLKNAVIVIGMLARPOSHPTDPPDKITTYVSCPNRNSLLENTSIDLISVYGFSAF 746
 Db 698 VQIYVMSVWMLLEPGRRTTLPEKRETVILKCNVK--DSGMLSLITVDDVYLILCYTAA 756
 QY 747 MGKELPNTYNAKFTTYSMTFTSSVSLCTFMSAISGVLTVYDLVLY-VLNLAIISLG 805
 Db 757 KTRKCPENFENAKFISFTM---YTTCIWMALFIPF---YVTSDDYVQVTTMCISSVLS 810
 QY 806 -----YFGKCYMLIEPERN 821
 Db 811 GFVVLGLCLFAKVAHITLPOPKN 833
 RESULT 9
 JC7160
 C:Species: Mus musculus (house mouse)
 C:Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 11-May-2000
 C:Accession: JC7160
 R:Minoshima, T.; Nakanishi, S.
 J. Biochem. 126, 889-896, 1999


```

0Y 29 LRGDVLGGGSLNHNKGIYHNLQYPMCKEYKVTGYNLMQAMFAVEIINBSSL 88
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 35 IEGDLVGLGGLPPINKGTGTGTEBCRGINEDR-----GIQRLEAMFLAIDEIKDYL 85
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 89 LRGVLLGGEIYDVC-----YISNNVQPVLYFLAHEDN---LLP-----IQEDSNYSIRV 135
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 86 LRGVYLGHIIIDPCSRDTYALQESLEFVRASLTJKVDEAEVCMPOGSAIQE--NIPLLI 142
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 136 VAVIGAPDMSSEMYVANTLSLELLPQITYSALISDELPRKYFPFALLRTTPSDHIVEMV 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 143 AGVIGGTSYSYIQVANTLRLFPQIPISTASTALSUKSRIDYFARKVDPDEFYQAKMA 202
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 196 QLMLEFRNNMIIIVYSSDTYGRNGQLIGERYARRDLICIAOETLPITQAPONNTSEERQ 255
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 203 ELIRFENMTYVSVASEGDTGTYEAFEGEARLNICIATAEAKVGRSNIKRSDS----- 258
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 256 RLVTLYVDLQOSTAVYVYVSPDLTLHFENEYLRQN--FTGAWIAIASENALDPVYHLN 313
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 259 ---VIRELLQKRNAYVVLVFMKSDSRELIAASRYNASF--WVASDGGQAOESITYKG 312
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 314 TELGHLEPTLITTIQSVPI-----PGFSEFRMGQAPPLSRTSQ 355
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 313 SEHVAYGA-IILELASHVYRQEDRFQSLNPNNNHRNWFDFWE-----Q 357
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 356 SYTCN-----QECNCLNATJSEFTILRLSGERVVYSYSAVYAVAHALHSLLG--C 405
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 358 KFOCSLQNRKRNHRQICDCHLAIDS-----NYEQESKIMFVYVNAVYAAHALLHKMQRTLC 412
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 406 DKST--CTKRVVYPMQL--EEIMKVNFTL-----LDHOIFPDPOGDVALHLEIYQO 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 413 PWTYTLCDAMKILDSKRLIKYDLKINTFTAPRPNPKGADSYKEDTYDGGKRIYVWFQ 472
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 455 WDRSONPQSOVASYYPRLQRLKNIQDISMHTVNTNITPMSKSKRCQSQGKKRPVGIHVC 514
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 473 HIGGYSYLYLKQHW--AETLYLDVDSIHMS--RNSVPTSOCSDPCAPRPMKRMPOGDVCC 528
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 515 FECIDCLGFTLNHDEDEYEOACQNNMWSYOSEISCKRQRLVLEHMEARTIYAVALLA 574
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 529 WLCICEPEYEL---VDEFTCDGCGPQMPTADISGCTYNLPEDYIRWEDANAIGEVYTIAC 585
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 575 LGLISTLAILFWNRHFOPIVYRSAGGPMCELMTLTLVAYAVVYVYVQPKRVSTCICRQ 634
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 586 LGMCTCIVITVEYFIKHNNMTPLYKASGRELCYILLGVSILSCMPPEFLAKSPYICALRR 645
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 635 ALFPLCEFTICISCIANVSFOIYCAF--KMASRPRAISYWRKQGP-----YVSARFTY 687
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 646 LGLGTSFAICSAIITLTKNCIARIPFDGVKNGAQRKFTS-----PSSQYFICGLIIV 698
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 688 LKMVTVYIGMLARPOSHRTPDDPKITIVSCNPYRNSLLEFNTSIDLLSVGGSFAYM 747

```

[illegible]

```

      RESUFI 10
JH0561
metabotropic glutamate receptor 2 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1992 #sequence Revision 30-Jun-1992 #text_change 13-Sep-1998
C:Accession: JH0561
R:Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
Neuron 8, 169-179, 1992
A:Title: A family of metabotropic glutamate receptors.
A:Reference number: JH0561; MUID:92110002; PMID:1309649
A:Accession: JH0561
A:Molecule type: mRNA
A:Residues: 1-872 <TAN>
A:Experimental source: brain
C:Comment: This protein is coupled to a G protein and evokes a variety of functions b
C:Superfamily: metabotropic glutamate receptor 4
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane p
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-872/Product: metabotropic glutamate receptor 2 #status predicted <MET>
F:568-590/Domain: transmembrane #status predicted <TRI>
F:605-625/Domain: transmembrane #status predicted <TRI>
F:637-655/Domain: transmembrane #status predicted <TIY>
F:680-700/Domain: transmembrane #status predicted <TIY>
F:726-747/Domain: transmembrane #status predicted <TRY>
F:761-783/Domain: transmembrane #status predicted <TVI>
F:795-819/Domain: transmembrane #status predicted <TVI>
F:803,286,338,402,547/Binding site: carboxylate (asn) (covalent) #status predicted
F:601,675,827,837,843/Binding site: phosphate (Thr) (covalent) #status predicted
F:832/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match          15.7%; Score 699.5; DB 2; Length 872;
Best Local Similarity 26.8%; Pred. No. 2.3e+45;
Matches 234; Conservative 157; Mismatches 365; Indels 117; Gaps 33;

QY   14 LLM-VLAERAEKSDYLRLGDVLLGGSLFLSH-----ANNKGIVHLNLFQVPCKEYEKVI 67
    |||  |||  :||  |  ||  ||||||  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|
DB   12 LLMGAVARSPAKKVLTLEGGDVLVGLFPVHQGPAEECGPN-----EHR----- 57
    |||  |||  :||  |  ||  ||||||  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|
QY   68 GYNLQAARFAVEELINDSSLPGVLLGYEIVDYCYISNN-VOPLYELAHE-----DNL 121
    |  :||  |||  :||  :|  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB   58 GIQRLEAMLFALDRINRDPRPHLLPVGRLGAHLIDSCSKDTHALDEQLDFRASLSRGADGS 117
    |||  |||  :||  |  ||  ||||||  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|
QY   122 LPIQED-----YSNTISKRYAVIGPDNSESWITANFSLSTLPQTYSALSDELQKYR 176
    |||  |||  :||  |  ||  ||||||  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|
DB   118 RHICPDGSYATHSDAPRTAVTGVIIGSGSYDSVTQVANILRLQIPDISAYASTARKLSDSR 177
    |||  |||  :||  |  ||  ||||||  :|  :|  :|  :|  :|  :|  :|  :|  :|  :|
QY   177 PfallRTTPSADHNIEAMVOLMLHFRRMMIIIVLVSSDPYGRDNCOGLGERVAARDICIAF 236
    ||  ||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
DB   178 VYEFARYVPPPDFOKAAAEILRFNMWTYYGVSVASBSGDIETAFELRARINICAT 237
    |||  |||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
QY   237 QETPLTLQPNOMNISSEERQRLTYTDKIQOSTARVVVF--SPDLITYHFFNEVLRONFT 294
    |||  |||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
DB   238 SEKV-----GRAMSRRAAEQGVVRAL-LQKPSARVAVLFTSEARELIATAORIMASF 290
    |||  |||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
QY   295 GAVWYIASNSM-AIDPVLLNIPLGHILGFILGITIOSVPIIPESSEER----WGPOAGPP 349
    |||  |||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
DB   291 ---WAASDGMALESVAVG-SERAEGA-TITLELASTYISTSPASTFOGLDPMNNSRNPF 345
    |||  |||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
QY   350 ISRTSQSYTCNOECNCINATLSFNTILRLSGERYVYSVAUYAVAHAHSL--LGCDK 407
    ||  ||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
DB   346 REFMEERFCSPRODCAAHSI---RAVPFESEKIMEFYVAAYAMAALINMRALCPN 402
    ||  ||  :||  :||  :||  :||  :||  :||  :||  :||  :||

```



```

Db      30 HMPGDIIGALFVSF-----H-----QPTVDKVERKCGAVREOYGQIORVAMLHTLER 78
      82 INNDSSLPGVLLGGEIYDVCYSN-NVQPVLYFLAHEDNLLPIQED-----YSN 130
      79 INSDPTLLPNTLLGCEIDSCWHSVALEQSTIEFT--RDSLISSEEGGLVRCVDGSSSS 136
      131 YISR--VVAVIGPDNSESMTVANFLSLFLPOLITYSAISDELBRKVFAPALLRTTPSAD 188
      137 FRSKRPYIGVIGPSSVAIQVQLQLENIPQIAVSATSMDSKTLFKYEMRVVPSDA 196
      189 HVEAMVOLMLHFRRNMWIIIVLSSDPTGRDNGQLGERVARDICIAFOELPTLOPNON 248
      197 QOARMAVDIVKRYNMTYSAVHTEGNYGESGMEAPKDSAKGICIAHSYKRYSNAGSOS 256
      249 MTSEERORLVTIVDKLOOS-----TARVYVSPDLTLHYEFNEVLKRONFGA-VWIAS 301
      257 F-----DKLLKLTSHLPKARVAVACFCGKMTVRGLLMMRRLLAGLAGEFLILGS 304
      302 ESWAIDPVHLNLTGLHGTFLGITI--QSVPI-----PGSEPRFW 341
      305 DQMA---DRYDVTGQYQREAVAGGITIKLOSPDKWFDYYLKLRETNHNRNWFQEFWQH 361
      342 GPQAGPPPLSRYSOSYTCNOECDNCLNATLSNTILRLSGERYYS---YSAVAYAAH 397
      363 RFQCRLEGPPOENSKY--NKTC-----NSSLTILKTHVODSKMGFVINALYSMAK 409
      398 ALHSLGCDKSTC-----TKRVYVPMOLLEELIMKNFT--LIDHQIFPDPOGVALHL 448
      410 GLHNM--QMSLCPGYAGLCDAKMRPIDGRKLLSELMKTNFTGVSQDTLLFDENGSPGRY 466
      449 EIVOMQ-----MDRSQNPQSVASVYPLOROLKNIDISWHTVNNTPISM 494
      467 EIMNFKEMGKDYFDYINGSMONG-----ELKMDDEWMSKSNII--RSV 510
      495 CSKRQSGO---KKRPVGIHVCCPECIDCLPRTFLNTEDEYECQACPNNMWSQSETS 550
      511 CSEPEKQIKVIRKGEVS---CWTCTPCKENEYV---FDEYTCACQOLGSMPTDILTG 564
      551 CEKROLVLEWHEAPTIIVALLAAGFLSTLAIIVIFMRHPTPIVRSAGGPMCLMTL 610
      565 CDLIIVQILRMGDPEPIAAVAVFACGLATLFTVYVFIIRPTPVKSSRRLCIIILAG 624
      611 LTVAYMVPVYVGPVKVSTCLCQALFPLCTICISCIASVRSFOIVCAF---FMASRF 665
      625 ICLGILCTFCLAKKQIYCYLQIRIGISPMASYSALVTKNRIARILAGSKKICRYK 684
      666 PRAYSYWRVYQGPVYSMAFTV-LKMWIVVIGMLARPOSHPTDDPKRITIVSCNPNYR 724
      685 PRFMSACAO---LVIAFILICIQIGITVAFIMEP-----PD-----IMHDYPSIR 727
      725 N-SLFLNTSLDLLLSVGFES-----FAYMGKELPTNYNEAFITLSMFTYTSVS 774
      728 EYLLICNTNTNGVVPPLGNGILLISCIFYAKTRNVANFEAKYIAFTM---YTTGCI 784
      775 ICTFMSATSGVLTIVDLLVTVNLAIISLG-YFGPKCMILFYPERNTPAYFNS 828
      785 WIAFPPIYFGSNYKIITIMCFSVLSATVAGCMFVPKYIILAKERNVRSAFIT 839

```

RESULT 13

metabotropic glutamate receptor 5 B - human
 C:Species: Homo sapiens (man)
 C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 05-Jan-1996
 C:Accession: JC2131
 R:Minakami, R.; Katsuki, F.; Yamamoto, T.; Nakamura, K.; Sugiyama, H.
 Biochem. Biophys. Res. Commun. 199, 1136-1143, 1994
 A:Title: Molecular cloning and the functional expression of two isoforms of human metabo
 A:Reference number: JC2131; MUID:94197696; PMID:7908515
 A:Accession: JC2131
 A:Molecule type: mRNA
 A:Residues: 1-1212 <MIN>

C:Comment: This protein is coupled to guanine nucleotide binding proteins.
 C:Keywords: glycoprotein; neurotransmitter; receptor; transmembrane protein
 F:580-604/Domain: transmembrane #status predicted <TM>
 F:617-637/Domain: transmembrane #status predicted <TM>
 F:647-664/Domain: transmembrane #status predicted <TM>
 F:694-714/Domain: transmembrane #status predicted <TM>
 F:738-759/Domain: transmembrane #status predicted <TM>
 F:773-784/Domain: transmembrane #status predicted <TM>
 F:803-827/Domain: transmembrane #status predicted <TM>

Query Match 15.2%; Score 675.5; DB 2; Length 1212;

Best Local Similarity 25.7%; Pred. No. 2,5e-43;

Matches 230; Conservative 153; Mismatches 333; Indels 179; Gaps 34;

```

      28 YLPGLYLLGLGFLSHANKGLVHLNFIQVPMCKEYEVY-----IGYNMQAMRFAVE 81
      30 HMPGDIIGALFVSF-----H-----QPTVDKVERKCGAVREOYGQIORVAMLHTLER 78
      82 INNDSSLPGVLLGGEIYDVCYSN-NVQPVLYFLAHEDNLLPIQED-----YSN 130
      79 INSDPTLLPNTLLGCEIDSCWHSVALEQSTIEFT--RDSLISSEEGGLVRCVDGSSSS 136
      131 YISR--VVAVIGPDNSESMTVANFLSLFLPOLITYSAISDELBRKVFAPALLRTTPSAD 188
      137 FRSKRPYIGVIGPSSVAIQVQLQLENIPQIAVSATSMDSKTLFKYEMRVVPSDA 196
      189 HVEAMVOLMLHFRRNMWIIIVLSSDPTGRDNGQLGERVARDICIAFOELPTLOPNON 248
      197 QOARMAVDIVKRYNMTYSAVHTEGNYGESGMEAPKDSAKGICIAHSYKRYSNAGSOS 256
      249 MTSEERORLVTIVDKLOOS-----TARVYVSPDLTLHYEFNEVLKRONFGA-VWIAS 301
      257 F-----DKLLKLTSHLPKARVAVACFCGKMTVRGLLMMRRLLAGLAGEFLILGS 304
      302 ESWAIDPVHLNLTGLHGTFLGITI--QSVPI-----PGSEPRFW 341
      305 DQMA---DRYDVTGQYQREAVAGGITIKLOSPDKWFDYYLKLRETNHNRNWFQEFWQH 361
      342 GPQAGPPPLSRYSOSYTCNOECDNCLNATLSNTILRLSGERYYS---YSAVAYAAH 397
      363 RFQCRLEGPPOENSKY--NKTC-----NSSLTILKTHVODSKMGFVINALYSMAK 409
      398 ALHSLGCDKSTC-----TKRVYVPMOLLEELIMKNFT--LIDHQIFPDPOGVALHL 448
      410 GLHNM--QMSLCPGYAGLCDAKMRPIDGRKLLSELMKTNFTGVSQDTLLFDENGSPGRY 466
      449 EIVOMQ-----MDRSQNPQSVASVYPLOROLKNIDISWHTVNNTPISM 494
      467 EIMNFKEMGKDYFDYINGSMONG-----ELKMDDEWMSKSNII--RSV 510
      495 CSKRQSGO---KKRPVGIHVCCPECIDCLPRTFLNTEDEYECQACPNNMWSQSETS 550
      511 CSEPEKQIKVIRKGEVS---CWTCTPCKENEYV---FDEYTCACQOLGSMPTDILTG 564
      551 CEKROLVLEWHEAPTIIVALLAAGFLSTLAIIVIFMRHPTPIVRSAGGPMCLMTL 610
      565 CDLIIVQILRMGDPEPIAAVAVFACGLATLFTVYVFIIRPTPVKSSRRLCIIILAG 624
      611 LTVAYMVPVYVGPVKVSTCLCQALFPLCTICISCIASVRSFOIVCAF---FMASRF 665
      625 ICLGILCTFCLAKKQIYCYLQIRIGISPMASYSALVTKNRIARILAGSKKICRYK 684
      666 PRAYSYWRVYQGPVYSMAFTV-LKMWIVVIGMLARPOSHPTDDPKRITIVSCNPNYR 724
      685 PRFMSACAO---LVIAFILICIQIGITVAFIMEP-----PD-----IMHDYPSIR 727
      725 N-SLFLNTSLDLLLSVGFES-----FAYMGKELPTNYNEAFITLSMFTYTSVS 774
      728 EYLLICNTNTNGVVPPLGNGILLISCIFYAKTRNVANFEAKYIAFTM---YTTGCI 784
      775 ICTFMSATSGVLTIVDLLVTVNLAIISLG-YFGPKCMILFYPERNTPAYFNS 828
      785 WIAFPPIYFGSNYKIITIMCFSVLSATVAGCMFVPKYIILAKERNVRSAFIT 839

```

RESULT 14

A46742 metabotropic glutamate receptor, mGluR6 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000

C:Accession: A46742

J. Nakajima, Y.; Iwakabe, H.; Akazawa, C.; Nawa, H.; Shigemoto, R.; Mizuno, N.; Nakanishi, S. *Biol. Chem.* 268, 11868-11873, 1993

A:Title: Molecular characterization of a novel retinal metabotropic glutamate receptor

A:Reference number: A46742; MUID:93280152; PMID:8389366

A:Accession: A46742

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-871 <NAK>

A:Cross-references: GB:D13963; NID:g391856; PIDN:BA03066.1; PID:g391857

A:Experimental source: retina

A>Note: sequence extracted from NCBI backbone (NCBIN:133246, NCBIPI:133250)

C:Superfamily: metabotropic glutamate receptor 4

C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 15.1%; Score 673; DB 2; Length 871;

Best Local Similarity 26.4%; Pred. No. 2,5e-43; Matches 239; Conservative 155; Mismatches 377; Indels 134; Gaps 33;

Db 11 LEFLWLAE--DAENSDFTLPDGYLLGLFSLHANNKGIHLNLTQVPMCKEYEVKI 67

Db 8 LMLMLWLSQAGIAGAGSVLAGLTLGLFPAHGAAGACGALK-----KEQ 58

QY 68 GYNLMQMRFAVEIINDSSLLPVLGLEYLVDCYISNNVQPVLYPLAHEDNLLPIQED 127

Db 59 GVHLEMLALYLDVNMADPELLPVGRLGARLLDTC--SRDYALQALSFQALIRSGD 116

QY 128 -----YSNYISRVVAIVGPDNSESVMVTANFLSFLPLQITVYASIDELR 172

Db 117 GDEASVRCGCVPLRSPRRVAVVAGASSVISVAVANLRFALPQISTASTABELS 176

QY 173 DKVFPALLRTTPSADHHEAMVOLMLHFRWNWIIIVSSDPTGRDNGQLGEEVAR-- 230

Db 177 DSTRYDEFSRVVPPDSYQOAMVDIVALGMVYSTLASSENG-ESGVFAVQISREAG 235

QY 231 DICIAPQETLPTLPQNMNTSEERQVLYIDKIQOSTARVVVPSDLTIHFENEYLR 290

Db 236 GVCIAQOSTIKIPR-EPKGEERKIVRL-----METPNARIIIFAMEDDIRVLEATRQ 288

QY 291 QNFTG-AVMIASESMA--IDVLIHLNTE--LGHITFLGII--QSVPIPFSEFRWGP 343

Db 289 ANLGHFLWGSDSWGSKISPII-LLEEAVG-----AIIILPKRASIDGFDQYF-- 337

QY 344 QAGPPILSRFSQSYTCN-----QECDCNCL--LNTLSFNTIILRSGER----- 383

Db 338 -----TTRSLNNRRNRINWFAEFEEENFNCKLTSSGGOSDSTFKCKEERIGODSAVEQ 391

QY 384 --VYYSVYSAVVAHAHLHL--LGCDKSTCKTRVYVPMQ--LLEETWVKNFT-LLDHQ 435

Db 392 EGKQVQYIDAVYATAHLSHQAQLCPGHGLCPAMEPTDGRITLHYTRAENGSGATP 451

QY 436 IFPDPOGVALHLEIYVOM--MDRSONPQSVASYRPLQROLKNIDISWHTYNTNTPM 492

Db 452 VMFENDADAGRYDIFQYQANTNSASSGGYQAVGOMAEALR--LDMEVLRSGDPHEVPP 509

QY 493 SMCKRCOSGOKKRPVGIHVCCFECIDCLPQTLNHTEDYECQACPNENMSYQSETSCF 552

Db 510 SCSLPCGPERKKRVKGVPCMCCEACDGYRF--QVDEPTCEACGDMRPTNHTGCR 566

QY 553 KRQVLEFHEHAPPIAVALLAAGLSTALIVFMRHPTPIYRSAGGPMCFMLTL 612

Db 567 PTPVVRILTWSSPMALPCLLAVLIGIMATYTIMATFMRHNDPIYRASGRELSYVLTLGIF 626

QY 613 VAVWVVVVVYVPPKVTCLQALFPLCFITICICIAVRSPQIYCAKMASTRFRAVSVM 672

Db 627 LITAITFLVMAEPCALICARRLILGIGTILSLITLTKNRIYRIFEGCKRSVT----- 681

QY 673 VRYQPYVS-----MAF-ITVLKMYIVVIGMLARP-----QSHPRTPDDPKITIVS 718

Db 682 ---PPFISPTSQLVIRFGILSLQVGVIVAMIGAQPHSVIXDVEQRTVDPQEAR-GVLR 737

QY 719 CNPNYRNSLFPNMSLDLLSVGSEFAYMCKELPTNNAKFTTSLATFTYSSVSLCTP 778

Db 738 CDMS-DLSLIGCLGCLSLLLVTCIVVAIKARGVPEFENEKPIGFV--YTCIILWAF 793

QY 779 MSATSG-----VLVYIVDLVTVLNLALISIG-YRGPCCYMLFPERRTPAFENSMI 830

Db 794 VPVFPGIAQSAEKIYIGTITLVVLSLSASVSLGMLYVPTTYILFHPEDONQRRRSKL 853

QY 831 QGYTM 835

Db 854 KTSFM 858

RESULT 15

I49142

metabotropic glutamate receptor 8 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999

C:Accession: I49142

R. Duvoisin, R.M.; Zhang, C.; Ramonell, K.

J. Neurosci. 15, 3075-3083, 1995

A:Title: A novel metabotropic glutamate receptor expressed in the retina and olfactory

A:Reference number: I49142; MUID:95239344; PMID:7722646

A:Accession: I49142

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-908 <RES>

A:Cross-references: EMBL:U17252; NID:g854728; PIDN:AAA68149.1; PID:g854729

C:Genetics:

A:Gene: mGluR

C:Superfamily: metabotropic glutamate receptor 4

C:Keywords: neurotransmitter receptor

Query Match 15.1%; Score 672; DB 2; Length 908;

Best Local Similarity 25.2%; Pred. No. 3.1e-43; Matches 230; Conservative 176; Mismatches 375; Indels 132; Gaps 37;

QY 4 RAKTICSLFLL-----WLA--EPANSDP-----YLPDYLGLGFLSLA-NMKGIYHL 51

Db 6 KRSTSCPCFELIAKFPWILTMQRTSDEYASIRLDGIIIGGLFPAHAKERGCV-- 62

QY 52 NFLQVPMCKEYEVKIVGYNLMQMRFAVEIINDSSLLPVLGLEYLVDCYISNNVQPV 111

Db 63 -----PCGDLK-KEKIHLEMLYAIIDQTNMDPDLSSITITGLVRLDTC--SRDYAL 113

QY 112 LVFLAHEDNLLPIQEDYSNT-----ISRVVAIVGPDNSESVMVTANFLSFL 159

Db 114 EQSLTFVQAL--TEKQASDVKACANGDPIFTKPKDKISVIGAAASVSIVAVANLRLFKI 171

QY 160 PQTYSASIDELNDKVFRLRTTPSADHHEAMVOLMLHFRWNWIIIVSSDPTGRDNG 219

Db 172 PQTYSASTAPLSDNTRYDEFSRVVPPDSYQOAMVDIVALGMVYSTLASSENG-ES 230

QY 220 GOLLEGEVAR--DICIAPQETLPTLPQNMNTSEERQVLYIDKIQOSTARVVVPS 277

Db 231 GVAFQIISIEIGVCAQSQKIPR-EPKGEERKIKRL-----LETPNAARVIMFAN 283

QY 278 DLTLYHFENLEVRONTG-AVMIASESMA--IDVLIHLNTELGHTFLGII--QSVPI 332

Db 284 EDIRGLLEAKKLNGSGHFLWIGSDWSGKIAPV-YQGEIAE-----GAVTILPKRASI 338

QY 333 PG-----FSEPRE--WGPQAGPPPLSTFSQSYTCNQCDCNCLNATLS 372

Db 339 DGFDRYRSRTLANNRNRMVFAFSENGFCCKSGSHG-KRNSIRKCT-----G 386

QY 373 FNTILRSQ---ERVYYSVSAVVAHAHLHLSGDQKSTCRV-----VYPMQL 421

Db 387 LERTARDSYEQBQKQVIVDAVYSMAVYALHNM--HKEICPGYIGLCPRMTIIDGKELL 443

OY 422 BEIWKVNT-LIHDQIFEDPOGDVALHLEIYOMQMDRSONPROQSVASYYPLOROLKNIOD 480
Db 444 GYIRAVNNGSAGTPTVFNENGADGRYDIFQYQINNKSTYK-TIGHWTQNLHLK-VED 501
OY 481 ISMHTVNTTIPMSCKRCSOGOKKRPVGIHVCCFECIDCLPGTFPLNTEDEYECQACPN 540
Db 502 MQMANREHTHPASVCSLPCFKGERKKYKGVPCCHGRC---EGYNYQVDELSCELCPL 558
OY 541 NEMSQSTSCFKROLVLEHHEAFTIYVALLAALGFLSTALIVIFWNRHROTPIYRSAG 600
Db 559 DQRPINRTGCRPIIKLEHSPWAVVPVLAIALIGIATTEVIYTFVRYNDTPIYRASG 618
OY 601 GPMCFIMLTLLVAVMVPVYVGPVKVSTCICROALFPLCFYICISCIJAVRSFOIYCAFK 660
Db 619 RELSTVLTGIFLCYSTIFLMIAPDPTIICSPRRIFLGLMCFSYAALLTKNRHRIPE 678
OY 661 MASRPPRAYSYWRYQGPVSNMFTVLKMYIVVIGMLARQOS---HPRT-DEDDPKI 714
Db 679 QGKKSVPAPKFISPASQLVITFSLISYQLGVEFWFVVDPPHTIIDYGEQRTLDPENAR- 737
OY 715 TIVSCNPNYRNSLENTSLDLLLSYGFSEFAYMGKELPTNYNEAKFTLSMTFTSSYS 774
Db 738 GVLKCDIS-DLSLCSLSIGSILMWTCTVYAIKTRGVDETFNEAKPIGFTM---YTCII 793
OY 775 LCTFMSAYS-----VLVTIVDLVTVLNLALISLG-YFGPKCYMILEYPERNT---P 823
Db 794 WIAFTIPFPGTAQSAEKMYIQTTTLTVSMLSASVSLGMLWPKYIIIIHFPEQNVQKRK 853
OY 824 AYFNSMIGQYTW 836
Db 854 RSFKAVYTAATMQ 866

Search completed: May 19, 2003, 09:51:13
Job time : 24.8343 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:46:50 ; Search time 10.9089 Seconds
(without alignments)
3186.136 Million cell updates/sec

Title: US-09-927-315-9
Perfect score: 4443
Sequence: 1 MGRPRATICSLFILWVLAEE.....ERNTPATFNSMIGYTRMRD 838

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt.40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1094	24.6	1078	1 CASR_HUMAN	P41180 homo sapien
2	1088.5	24.5	1079	1 CASR_RAT	P48442 rattus norv
3	1087.5	24.5	1085	1 CASR_BOVIN	P35384 bos taurus
4	1085.5	24.4	1079	1 CASR_MOUSE	O9GY96 mus musculu
5	724.5	16.3	915	1 MGR7_RAT	P35400 rattus norv
6	721.5	16.2	912	1 MGR4_RAT	P31423 rattus norv
7	721.5	16.2	915	1 MGR7_HUMAN	Q14831 homo sapien
8	718	16.2	912	1 MGR3_RAT	P31422 rattus norv
9	717.5	16.1	912	1 MGR4_HUMAN	Q14833 homo sapien
10	709	16.0	877	1 MGR3_HUMAN	Q14832 homo sapien
11	699.5	15.7	872	1 MGR2_RAT	P31421 rattus norv
12	689	15.5	976	1 MGR_DROME	P31685 drosophila
13	686.5	15.5	872	1 MGR2_HUMAN	Q14416 homo sapien
14	685	15.4	1203	1 MGR5_RAT	P31424 rattus norv
15	675.5	15.2	1212	1 MGR5_HUMAN	P41594 homo sapien
16	674.5	15.2	877	1 MGR6_HUMAN	O15303 homo sapien
17	673	15.1	871	1 MGR6_RAT	P35349 rattus norv
18	673	15.1	908	1 MGR8_RAT	P70579 rattus norv
19	672	15.1	908	1 MGR8_MOUSE	P47743 mus musculu
20	666.5	15.0	908	1 MGR8_HUMAN	O00222 homo sapien
21	655.5	14.8	999	1 MGR1_CAELI	O09630 caenorhabdi
22	654	14.7	1199	1 MGR1_RAT	P23385 rattus norv
23	647	14.6	1194	1 MGR1_HUMAN	Q13255 homo sapien
24	180.5	4.1	960	1 MGR1_MOUSE	Q94W18 mus musculu
25	177.5	4.0	961	1 MGR1_HUMAN	Q94B55 homo sapien
26	169	3.8	991	1 MGR1_RAT	Q92044 rattus norv
27	147.5	3.3	940	1 MGR2_RAT	O88871 rattus norv
28	144	3.2	918	1 MGR1_HUMAN	P39086 homo sapien
29	144	3.2	949	1 MGR1_MOUSE	P22756 rattus norv
30	140.5	3.2	836	1 MGR1_MOUSE	O60934 mus musculu
31	139.5	3.1	941	1 MGR2_HUMAN	O75899 homo sapien
32	131.5	3.0	1125	1 MGR1_MOUSE	P16065 strongyloce
33	130.5	2.9	938	1 NM2L_HUMAN	Q05586 homo sapien

ALIGNMENTS

34	129.5	2.9	938	1	NM2L_RAT	P35439 rattus norv
35	126.5	2.8	938	1	NM2L_MOUSE	P35438 mus musculu
36	125.5	2.8	683	1	YR86_CAELI	P34352 caenorhabdi
37	123.5	2.8	501	1	P325_HUMAN	Q94126 homo sapien
38	120.5	2.7	907	1	GLR1_RAT	P19490 rattus norv
39	118.5	2.7	907	1	GLR1_MOUSE	P23818 mus musculu
40	118	2.7	893	1	BOSS_DROI	O24738 drosophila
41	116	2.6	359	1	AG2R_RABIT	P34976 oryctolagus
42	116	2.6	4367	1	DTHC_NEUCR	P45443 neurospora
43	114.5	2.6	1047	1	ANPB_HUMAN	P20594 homo sapien
44	114.5	2.6	1108	1	CYGF_HUMAN	P51841 homo sapien
45	114	2.6	359	1	AG2R_PIG	P30555 sus scrofa

RESULT 1

ID	CASR_HUMAN	STANDARD:	PRT: 1078 AA.
AC	P41180: Q13912; Q16379; Q16108; Q16109; Q16110;		
DT	01-FEB-1995 (Rel. 31, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Extracellular calcium-sensing receptor precursor (Casr) (Parathyroid Cell calcium-sensing receptor).		
GN	CASR OR GPRC2A OR PCARL.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RP	SEQUENCE FROM N.A.		
RA	Pearce S.H.S., Thakker R.V.;		
RL	Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Parathyroid;		
RX	MEDLINE=95279439; PubMed=7759551;		
RA	Garrett J.E., Capuano I.V., Hammerland L.G., Hung B.C., Brown E.M., Hebert S.C., Nemeth E.F., Fuller F.;		
RT	"Molecular cloning and functional expression of human parathyroid calcium receptor cDNAs.";		
RL	J. Biol. Chem. 270:12919-12925(1995).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Kidney;		
RX	MEDLINE=95408281; PubMed=7677761;		
RA	Aida K., Koishi S., Tawata M., Onaya T.;		
RT	"Molecular cloning of a putative Ca(2+)-sensing receptor cDNA from human kidney.";		
RL	Biochem. Biophys. Res. Commun. 214:524-529(1995).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=96343808; PubMed=8756555;		
RA	Fretschel M., Zink-Lorenz A., Hollloschi A., Hafner M., Flockert V., Rade F.;		
RT	"Expression of a calcium-sensing receptor in a human medullary thyroid carcinoma cell line and its contribution to calcitonin secretion.";		
RL	Endocrinology 137:3842-3848(1996).		
RN	[5]		
RP	SEQUENCE OF 643-908 FROM N.A.		
RX	MEDLINE=96193893; PubMed=8613532;		
RA	Bikle D.D., Ratnam A., Mauro T., Harris J., Pillai S.;		
RT	"Changes in calcium responsiveness and handling during keratinocyte differentiation. Potential role of the calcium receptor.";		
RL	J. Clin. Invest. 97:1085-1093(1996).		
RN	[6]		
RP	VARIANTS FHH GLN-185; LYS-297 AND TRP-795.		
RX	MEDLINE=94094324; PubMed=7916660;		
RA	Pollak M.R., Brown E.M., Chou Y.-H.W., Hebert S.C., Marx S.J., Steinmann B., Levi T., Seidman C.E., Seidman J.G.;		

RT "Mutations in the human Ca(2+)-sensing receptor gene cause familial
RT hypocalcemic hypercalcemia and neonatal severe
RT hyperparathyroidism.";
RL Cell 75:1297-1303(1993).
RN [7]
RP VARIANT ADH ALA-127.
RX MEDLINE=951179179; PubMed=7874174;
RX Pollak M.R., Brown E.M., Estep H.L., McLaine P.N., Klor O., Park J.,
RA Hebert S.C., Seldman C.E., Seldman J.G.;
RT "Autosomal dominant hypocalcemia caused by a Ca(2+)-sensing receptor
RT gene mutation.";
RT Nat. Genet. 8:303-307(1994).
RN [8]
RP VARIANTS FHH MET-62; CYS-66; MET-138; GLU-143 AND GLN-227.
RX MEDLINE=95243222; PubMed=7726161.
RX Chou Y.-H.W., Pollak M.R., Brandt M.L., Toss G., Arrqvist H.,
RA Atkinson A.B., Papapoulos S.E., Marx S., Brown E.M., Seldman J.G.,
RA Seldman C.E.;
RT "Mutations in the human Ca(2+)-sensing-receptor gene that cause
RT familial hypocalcemic hypercalcemia.";
RT Am. J. Hum. Genet. 56:1075-1079(1995).
RN [9]
RP SEQUENCE OF 1-61 FROM N.A. AND VARIANT FHH ALA-39.
RX MEDLINE=95403641; PubMed=7673400.
RX Aida K., Koishi S., Inoue M., Nakazato M., Tawata M., Onaya T.;
RT "Familial hypocalcemic hypercalcemia associated with mutation in the
RT human Ca(2+)-sensing receptor gene.";
RT J. Clin. Endocrinol. Metab. 80:2594-2598(1995).
RN [10]
RP VARIANTS NSHPT LEU-227 AND TYR-598.
RX MEDLINE=96292293; PubMed=8673635.
RX Pearce S.H.S., Trump D., Wooding C., Besser G.M., Chew S.L.,
RA Grant D.B., Heath D.A., Hughes I.A., Paterson C.R., Whyte M.P.,
RA Thakker R.V.;
RT "Calcium-sensing receptor mutations in familial benign hypercalcemia
RT and neonatal hyperparathyroidism.";
RT J. Clin. Invest. 96:2683-2692(1995).
RN [11]
RP VARIANTS ADHP THR-116; HIS-681 AND SER-806, AND VARIANT SER-851.
RX MEDLINE=96311554; PubMed=8733126;
RX Baron J., Winer K.K., Yanovski J.A., Cunningham A.W., Laue L.,
RA Zimmerman D., Cutler G.B. Jr.;
RT "Mutations in the Ca(2+)-sensing receptor gene cause autosomal
RT dominant and sporadic hyperparathyroidism.";
RL Hum. Mol. Genet. 5:601-606(1996).
RN [12]
RP VARIANT FHH ARG-174.
RX MEDLINE=97442275; PubMed=9298824;
RX Ward B.K., Stuckey B.G.A., Gutteridge D.H., Laing N.G., Pullan P.T.,
RA Rajczak T.;
RT "A novel mutation (L174R) in the Ca2+-sensing receptor gene
RT associated with familial hypocalcemic hypercalcemia.";
RL Hum. Mutat. 10:233-235(1997).
RN [13]
RP VARIANT FHH GLU-557.
RX MEDLINE=21603857; PubMed=11762693;
RX Nakayama T., Minato M., Nakagawa M., Soma M., Tohe H., Aoi N.,
RA Kosuge K., Seto M., Ozawa T., Kamatsuse K., Kokubun S.;
RT "A novel mutation in Ca2+-sensing receptor gene in familial
RT hypocalcemic hypercalcemia.";
RL Endocrine 15:277-282(2001).
CC -1- FUNCTION: SENSE CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: FOUND IN KIDNEY, BUT NOT IN BRAIN, LUNG,
CC LIVER, HEART, SKELETAL MUSCLE, OR PLACENTA.
CC -1- DISEASE: DEFECTS IN CASR ARE A CAUSE OF FAMILIAL HYPOCALCEMIC
CC HYPERCALCEMIA (FHH) AND NEONATAL SEVERE HYPERPARATHYROIDISM
CC (NSHPT) TWO INHERITED CONDITIONS CHARACTERIZED BY ALTERED CALCIUM

CC	HOMEOSTATIS. THE MUTATIONS REDUCE THE ACTIVITY OF THE RECEPTOR.
CC	FHH AFFECTED INDIVIDUALS EXHIBIT MILD OR MODEST HYPOCALCEMIA,
CC	RELATIVE HYPOCALCIOTRIA, AND INAPPROPRIATELY NORMAL PTH LEVELS. IN
CC	CONTRAST, NSHPT IS A RARE AUTOSOMAL RECESSIVE LIFE-THREATENING
CC	DISORDER CHARACTERIZED BY VERY HIGH SERUM CALCIUM CONCENTRATIONS,
CC	INSTANTIAL DEMINERALIZATION, AND PARATHYROID HYPERTROPHY. IN SOME
CC	CASES NSHPT HAS BEEN DEMONSTRATED TO BE THE HOMozyGOUS FORM OF
CC	FHH.
CC	- I - DISEASE: DEFECTS IN CASR ARE ALSO THE CAUSE OF AUTOSOMAL DOMINANT
CC	HYPOCALCEMIA (ADH) IN WHICH THE RECEPTOR IS ACTIVATED AT SUBNORMAL
CC	CAL(2+) LEVELS.
CC	- I - DISEASE: DEFECTS IN CASR ARE ALSO THE CAUSE OF AUTOSOMAL DOMINANT
CC	HYPOPARATHYROIDISM (ADHP). ADHP IS CHARACTERIZED BY HYPOCALCEMIA
CC	AND HYPERPHOSPHATEMIA DUE TO INADEQUATE SECRETION OF PARATHYROID
CC	HORMONE. SYMPTOMS ARE SEIZURES, TETANY AND CRAMPS.
CC	- I - SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sdb.ch/announce/)
CC	or send an email to license@isb-sdb.ch .
CC	-----
DR	EMBL; X81086; CAA56990.1; -;
DR	EMBL; U20759; AAA86503.1; -;
DR	EMBL; U20760; AAA86504.1; -;
DR	EMBL; D50855; BAA09453.1; -;
DR	EMBL; S83176; AAB46873.1; -;
DR	EMBL; S79217; AAB35262.2; -;
DR	EMBL; S68032; AAB29413.2; ALT_SEQ.
DR	EMBL; S68033; AAB29414.1; -;
DR	EMBL; S68036; AAB29415.1; -;
DR	EMBL; S81755; AAD14370.1; -;
DR	Genew: HGNC:1514; CASR.
DR	MIM; 601189; -;
DR	MIM; 145860; -;
DR	MIM; 239200; -;
DR	MIM; 601198; -;
DR	InterPro: IPRO01828; ANF_receptor.
DR	InterPro: IPRO00337; GPCR_Mgr.
DR	Pfam: PF00003; 7tm_3; 1.
DR	Pfam: PF01094; ANF_receptor; 1.
DR	PRINTS: PRO0248; GPCRMGR.
DR	PROSITE: PS00979; G_PROTEIN_REC_P.F3.1; 1.
DR	PROSITE: PS00980; G_PROTEIN_REC_P.F3.2; 1.
DR	PROSITE: PS00981; G_PROTEIN_REC_P.F3.3; 1.
DR	PROSITE: PS02059; G_PROTEIN_REC_P.F3.4; 1.
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; signal; Disease mutation; Alternative splicing; Polymorphism.
FT	SIGNAL 1
FT	CHAIN 20 1078
FT	DOMAIN 20 612
FT	TRANSMEM 613 635
FT	DOMAIN 636 649
FT	TRANSMEM 650 670
FT	DOMAIN 671 681
FT	TRANSMEM 682 700
FT	DOMAIN 701 724
FT	TRANSMEM 725 745
FT	DOMAIN 746 769
FT	TRANSMEM 770 792
FT	DOMAIN 793 805
FT	TRANSMEM 806 828
FT	DOMAIN 829 836
FT	TRANSMEM 837 862
FT	DOMAIN 863 1078
FT	CARBOHD 90 90
FT	CARBOHD 130 130
FT	CARBOHD 261 261
FT	CARBOHD 287 287
FT	CARBOHD 386 386

FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 595 595 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1085 AA; 121170 MW; 5D66B8C9CD13E47 CRC64;

Query Match 24.5%; Score 1087.5; DB 1; Length 1085;
 Best Local Similarity 31.6%; Pred. No. 9e-73;
 Matches 279; Conservative 161; Mismatches 347; Indels 97; Gaps 28;

11 LEFLAVLAEPANSDPFLPBDVLGLGLSLAHNMKGIVHLNLOVP---MCKEYEVXY 67
 11 LAFSTCTSAVGPDAQAKGKIDILLGLFPIHGV-AVADQDKSPRESVECTRYNR-- 67
 68 GYNLMQAMFAVEINNNDSILPGLVLYGEIYDVC-YISNNQPVLYFLA-HEDNLLPIQ 125
 68 GFMWQAMFAVEINNNDSILPGLVLYGEIYDVC-YISNNQPVLYFLA-HEDNLLPIQ 127
 126 E--DYSNYSIRVAVIGPDSSEVTVANFLSLFLPQITYSASIDELNDKVPALLRT 183
 128 EFCNCSEHIPSTIAVVGAGSGISTAVANLGLFYIPQVSASSRSLSNKQKSFLLRT 187
 184 TFSADHVMAMQOLMHEFMMNITIVSSDTYGRDNGQLGEPVARDICIAQETLPL 243
 188 IPDEHOATAMADILEYEFMMNVTIAADDDGPGIEKFRERAEERDIDISELI--- 244
 244 QPNQNTSEEROLTVITKLOOSTARVVVSPDLTYHFEVNEVRONPTGAVWIASHS 303
 245 ----SQSYSE-EKIQOVVEVIONSTAKVIVVSSGDLPLKEIYRNITGTIMIASER 299
 304 WAIIDPVNLHMLTELGH-LGTFLLGTTIQSVPIPGFSEF-----REMGAQA-- 345
 300 MASSSLI-AMPEYFHVVGCTIGFGLKAGQIPGRREFLQVHPKSVHNGFAKFEWETPN 358
 346 -----GPPPL-----SRTSOSYF-----CQNEQNCNANLSEFTIIRLSG 381
 359 CHLOEGAKPPLPDTFLKHEBEGGAKLSNPFAPRLTGE-ENISSVETPYMDYTHL-- 415
 382 ERVVSYSVAVVAHAVALSHSLGC-----DKSTCFKRVVPMOLLEIMKVFNTL-L 432
 416 -RSTYVAVLYVSIANAHDITCIPGRGLFNGSCADIKKVAWQVLYKHLRLNTSMN 474
 433 DHQIFDPGQDVALHEIYQWOMKQSNP--FQSVASYRPL-----OROLKNIQDISMHT 485
 475 GROYTFDECGDLAGNYSIINMHLSPEDGSIYFEKV--GYNNVYAKKGERLFINDEKILMSG 533
 486 VNNTIPMSCKRCSGOKKKPV-GIHVCCFECIDLPCTFLNHTDEHECCACPNMNS 544
 534 FSEVYFSCSRDCLAGTKKGLIEGPTCCFCEVCEPDEYSDET-DASACDRCPPDEWS 592
 545 YQSETSFCKROLVLEHMEAPTIYAVALLAAGLSTLALIVFMRHQPPIYVSAGGPMC 604
 593 NEHNTSCIAKEIEFLSWEPFGALFLFAVLGIFLAFVIGVIRKRNPIYKATRELS 652
 605 FLMALLLVAVVYVYVGPVYVSTCLCROALPFLCTTICISCIASVQIYCAF--KMA 662
 653 YLLSLFLOCSSSLFIEPQDWTICRLRQPAFGISFVLCIILKYNVLLVFEAKPI 712
 663 SREPRASYVWVRQGVYVMAFITYVKMIVYIGMLARQSHRPDRDPDKTIYVSCNN 722
 713 TSEHRR--WVGINTOELFLVCTEQIYCAIMLTAPSSRYNHELDEILFIYCHG 769
 723 YNSLLENTSLDLILSVGFSEFAMKELPTYNNEAKFTLSTEFSTSVSICPTMSAY 782
 770 SLMAAGELIGYCLLAALICFFAFKSRKLPENNEAKFTFSLFIETWIS---PIPAY 826
 783 SGVLTYIVDLLVYVNLALISIG---YGPCKYMLTFPERNT 822
 827 ASTYGFVS-AVEYIAIILASFGLLACIFENKYYIILFRPSHT 869

RESULT 4
 CASR_MOUSE STANDARD: PRT; 1079 AA.
 ID CASR_MOUSE

AC Q90Y96; 0089668; 088519; Q90Y95; Q90Z08; Q9R1D6; Q9R1Y2;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE Extracellular calcium-sensing receptor precursor (CAsR) (Parathyroid
 DE cell calcium-sensing receptor).
 GN CASR OR GPRC2A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
 RC STRAIN=C57BL/6; TISSUE=Kidney;
 RX MEDLINE=20092890; PubMed=10625662;
 RA Oda Y., Tu C.-L., Chang W., Ciumrline D., Koemueves L., Mauro T.,
 RA Elias P.M., Biale D.D.;
 RT "The calcium sensing receptor and its alternatively spliced form in
 RT murine epidermal differentiation.";
 RL J. Biol. Chem. 275:1183-1190(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM A), AND TISSUE SPECIFICITY.
 RC STRAIN=Black Swiss X 129/SV; TISSUE=Kidney;
 RX MEDLINE=20119279; PubMed=10652312;
 RA P1 M., Garner S.C., Flannery P., Spurney R.F., Quarles L.D.;
 RT "Sensing of extracellular cations in Casr-deficient osteoblasts.
 RT Evidence for a novel cation-sensing mechanism.";
 RL J. Biol. Chem. 275:3256-3263(2000).
 RN [3]
 RP SEQUENCE OF 256-600 FROM N.A. (ISOFORM A), AND TISSUE SPECIFICITY.
 RC TISSUE=Epiphyseal cartilage;
 RX MEDLINE=20043955; PubMed=10579354;
 RA Chang W., Tu C., Chen T.-H., Komuves L., Oda Y., Pratt S.A.,
 RA Miller S., Shoback D.;
 RT "Expression and signal transduction of calcium-sensing receptors in
 RT cartilage and bone.";
 RL Endocrinology 140:5883-5893(1999).
 RN [4]
 RP SEQUENCE OF 507-582 FROM N.A. (ISOFORM A).
 RC STRAIN=MMR; TISSUE=Brain;
 RA Hildebrand J., Ammon H.P.T., Wahl M.A.;
 RT Submitted (May-1998) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 562-814 FROM N.A.
 RC TISSUE=Kidney;
 RA Moawad T.T., Ricciardi D.;
 RT Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 646-799 FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=97231187; PubMed=9076582;
 RA Quarles L.D., Hartle J.E. II, Siddhanti S.R., Guo R., Hinson T.K.;
 RT "A distinct cation-sensing mechanism in MC3T3-E1 osteoblasts
 RT functionally related to the calcium receptor.";
 RL J. Bone Miner. Res. 12:393-402(1997).
 CC -1- FUNCTION: SENSES CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
 CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EPIDERMIS, KIDNEY AND CARTILAGE.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL; AF110178; AAD28371.1; -

DR EMBL: AF110179; AAD28372.1; -
 DR EMBL: AF128842; AAD40638.1; -
 DR EMBL: AF068900; AAC19388.1; -
 DR EMBL: AB027140; BAA77688.1; -
 DR EMBL: AF002015; AAC53252.1; -
 DR EMBL: AF159565; AAF00193.1; -
 DR MGI: MGI:1351351; Gprc2a.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm.3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PRO0248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_REC_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_REC_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_REC_F3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_REC_F3_4; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Alternative splicing.
 FT SIGNAL 1 19
 FT CHAIN 20 1079
 FT DOMAIN 20 612
 FT TRANSMEM 613 635
 FT DOMAIN 636 649
 FT TRANSMEM 650 670
 FT DOMAIN 671 681
 FT TRANSMEM 682 700
 FT DOMAIN 701 724
 FT TRANSMEM 725 745
 FT DOMAIN 746 769
 FT TRANSMEM 770 792
 FT DOMAIN 793 805
 FT TRANSMEM 806 828
 FT DOMAIN 829 836
 FT TRANSMEM 837 862
 FT DOMAIN 863 1079
 FT CARBOHYD 90 90
 FT CARBOHYD 130 130
 FT CARBOHYD 261 261
 FT CARBOHYD 287 287
 FT CARBOHYD 386 386
 FT CARBOHYD 446 446
 FT CARBOHYD 468 468
 FT CARBOHYD 488 488
 FT CARBOHYD 541 541
 FT CARBOHYD 594 594
 FT VASAPIC 461 537
 FT CONFLICT 45 45
 FT CONFLICT 304 304
 FT CONFLICT 410 410
 FT CONFLICT 566 566
 FT CONFLICT 595 595
 FT CONFLICT 610 610
 FT CONFLICT 814 814
 FT CONFLICT 889 889
 FT CONFLICT 906 909
 FT CONFLICT 1057 1057
 FT CONFLICT 1064 1064
 FT CONFLICT 1076 1076
 SO SEQUENCE 1079 AA; 120839 MW; AAF8BD472736D6E CRC64;
 Query Match 24.48; Score 1085.5; DB 1; Length 1079;
 Best Local Similarity 31.88; Pred. No. 1.3e-72;
 Matches 285; Conservative 162; Mismatches 329; Indels 121; Gaps 31;
 QY 10 SLFLLWVLAIPAEISDYLK-----GDVILGLFSLHNMGIYVLAFLQYPMCKEY 62
 DB 9 ALALATY-----HSSAYGPDQRAQKGGDIIIGLGFIRH---GVAAKD--QDIKSRPE 56
 QY 63 EVKVIYNL-----MOAMFAVEEINDSSLLPGVLLGEYIVDYC-YISNNVQPYLYPIA 116
 DB 57 SVECIIRNRFGRMLQAMIFALIEEINSPALLPMNTIGRFDCNVTSKALENTLSPIYA 116
 QY 117 -HEONLPIQIE--DYSNISRVAIVGIPDNSESVMYANFLSLFLLOITYSAISDELRL 173

DB 117 ONKIDSLNLEDFNCSESHISTIAVVGATGSGVSTAANILGLFYIPQVYASSSRLLSN 176
 QY 174 KVRFPALLRTPPSADHVEAMVQMLHFRMMWIIIVLSSDYYGDNQOLGGERARBDIC 233
 DB 177 KNQKSEPLRTIPNDHDATAADIIETFRMWWVTIAADDYGPGRGIEKREAEEDDIC 236
 QY 234 IAFQETPLTLPQNMNMTSEERORLVTVDKLOOSTARVVVSPDLTHYEFNEVLQNPF 293
 DB 237 IDFSLEI-----SOYSEEBEIQVVEYI---QNSTAVIYVFSSGPDLEPLIKETVRNI 288
 QY 294 TGAVWASEWADIPVLAHNLTELG-IGTFLGITIQSVPIPGSEF-----338
 DB 289 TGRIMLSEAWASSSLI-AMPEYHVVGRTIGFLAKAGIIPGFEPLQKVPKRVHNGF 347
 QY 339 -REMGPOA-----GPPLSRTSOSYCNQECNCAATLSFNTILRSGE-----382
 DB 348 AKERWEETFNCHLQDQAKGLPVDTFVRSH--EEGNRLNLSYAFRPL--CTGDEINIS 403
 QY 383 -----RVVYSYSAVAVAAHALSHLAG-----DKSTCKRVVYPMQLLEE 423
 DB 404 VETPYMGVEHRIISYNYIAVYSTAHALQDIYTCLPGRGLFTNSCADIKVEAMQYLKH 463
 QY 424 IMKVNFT-LLDHQLEFPDQGVALLHLEIVQMQRSONP--FQSVASYPL-----OROL 475
 DB 464 LRHLNFTNMGEQVTFPECGDLVGNYSILNMHLSPEDEGSIFYKEV--GYNNYANKGERLF 522
 QY 476 KNIDISWHVYNNITIPSMSCSKRCQSGOKKPV--GIHVCEPCIDCJPGFLNTEDEYE 534
 DB 523 INEKKILMSGSRVPPNSCRDQAGTRKGIIGEPCEVECECDGERSGGT--DASA 581
 QY 535 QACPNMWSYQSETSFQRQVLEWHEAPTAVALAALGFLSTALIVFMRHOTP 594
 DB 582 CDKCPDFFWSNENTYCIKAEIEFLAWTEPGLITFAVIGILFAVLGVFIKFRNTP 641
 QY 595 IYRAGCPMCIIMTILLVAMVYVVGPRKYSTCICRQALFPLCTICISCIAYVSQ 654
 DB 642 IVKATNRELSTLLFLSLCESSSLFIEGPQDTCRLRPAPFISVFLCISLIVTNR 701
 QY 655 IVCAF--KMASRFPRAVSYYVYQGPVYSAFIVLKMVIYVIMLARPQSHPTDDDP 712
 DB 702 VLAIFEAKEIPISFRK---WGLMQLQFLVPLCTFMQIVCIITLYLAPBSYNNHELD 758
 QY 713 KITVSCNPYRNSLNTLSL---DLLSVYGFSAVAGKELPYNNEANFTLSMTYF 769
 DB 759 EIRITFC---HEGSLMAGSLIGVTCILAAICFFFAFRSKRLPNFPAKRTIFESMLIF 815
 QY 770 TSSVSLCTFMAISGVLTVYDLYLVNLAISLG-----YFGKCYMIIFYPRNT 822
 DB 816 IWTIS---FIPAYASTYGRFVS-AVEVIAIILAAFGILACIFENKYVYIILFKPSRNT 868
 RESULT 5
 MGR7_RAT
 ID MGR7_RAT STANDARD; PRT; 915 AA.
 AC P35400;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 7 precursor (mgIur7).
 GN GRM7 OR GPRC1G OR MGUR7.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=94117433; PubMed=8285585;
 RA Okamoto N., Hori S., Akazawa C., Hayashi Y., Shigemoto R.,
 RA Mizuno N., Nakanishi S.;
 RT "Molecular characterization of a new metabotropic glutamate receptor
 mgIur7 coupled to inhibitory cyclic AMP signal transduction.";

RL J. Biol. Chem. 269:1231-1236(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE-Olfactory bulb;
 RX MEDLINE=94195260; PubMed=8145723;
 RA Saugstad J.A., Kinzie J.M., Mulvihill E.R., Segerson T.P.,
 RT Westbrook G.L.;
 RT "Cloning and expression of a new member of the L-2-amino-4-
 RT phosphonobutyric acid-sensitive class of metabotropic glutamate
 RT receptors.";
 RL Mol. Pharmacol. 45:367-372(1994).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 CC ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: WIDELY DISTRIBUTED THROUGHOUT THE BRAIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MG1UR4.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D16817; BA04092.1; -;
 DR EMBL: 006832; AAA2065.1; -;
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_RECPE_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECPE_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECPE_F3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_RECPE_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; glycoprotein; signal;
 KW Multigene family; Olfaction.
 FT SIGNAL 1 32
 FT CHAIN 1 32
 FT DOMAIN 33 915
 FT DOMAIN 33 590
 FT DOMAIN 591 615
 FT DOMAIN 616 627
 FT TRANSMEM 628 648
 FT TRANSMEM 649 654
 FT TRANSMEM 655 675
 FT DOMAIN 676 702
 FT TRANSMEM 703 723
 FT DOMAIN 724 753
 FT TRANSMEM 754 775
 FT DOMAIN 776 788
 FT TRANSMEM 789 810
 FT DOMAIN 811 825
 FT TRANSMEM 826 850
 FT DOMAIN 851 915
 FT CARBOHYD 98 98
 FT CARBOHYD 458 458
 FT CARBOHYD 486 486
 FT CARBOHYD 572 572
 SO SEQUENCE 915 AA; 102231 MW; F28AFC4C6454A6C2 CRC64;
 Query Match 16.3%; Score 724.5; DB 1: Length 915;
 Best Local Similarity 28.0%; Pred. No. 6.3e-46;
 Matches 257; Conservative 152; Mismatches 353; Indels 157; Gaps 41;
 QY 9 CSLFLLWLAEPANSDFTLP-----GDYLLGSLFSLHANKGIHLNLFQVPMCKEY 62
 DB 19 CVLEVLVLAALAAARGQEMKAPHSIRIEDVTLGGSLPFAHA--KGP-----SGVP-CGDI 70
 QY 63 EKVAVGYNLMQAMRAVEEINNDSLSLPGVLLGELIVDVYCISSNVQPYLYFLAHEDNITL 122

DB 71 K-RENGIRHLEAMLYALQDINSNPMLRPVTLGARILDTCC--SRDTYALEQSLTFVQAL- 126
 QY 123 PIQEDYSNY-----ISRVAVIGPDNSESVMTVANFLSLPQITYSASDE 170
 DB 127 -IQKDTSDVRCNTGEPPEVFKKPEKTVGVAGSGSSIVWANILRLFOIQTISYSTAPE 185
 QY 171 LBDKVRPALLETPTPSADHIVAMVQMLHFPMNNTIYVSSDTYGRDQGLSGREVAR 230
 DB 186 LSDRRITDFPSRVNPPDSFOAQAMVDYKALCMNVSTIASGSGTG-EGVESPTQISKE 244
 QY 231 --DICIARQETLPTLPQNPOMNTSEERQRLV---TIVDKQOSTARVVVFPSPDLTLVHF 284
 DB 245 AGGLCLQASVRP-----QERKDRITDPRILIKQLDIPNSRAVYTFANDEDIKQI 295
 QY 285 FNEVLRQNTG-AVNIASEMA--IDPVJHNTLGLHGLTFLGTLITQ--SVIIPGSEER 339
 DB 296 LAAARRADQVGFHFWGSDWSGSKINP-LHQBEDIAE--GAITQPRRAVEEGDAYF 350
 QY 340 EMGPQAGPPPLSRTSOSYCN-----QECDCN--LNATLSFMTILRSG-ERY- 384
 DB 351 -----TSRTLENNRNWVFAEYWEENCKLITSGSKEDTRKCTQGERIGKDS 400
 QY 385 -----YYSVSAVYAVAAHALSLGCDKSTCR-RVYTP-----WQLLEETKYNF 429
 DB 401 NYEQGKVQFVIDAVYAMAHALHHM--NKDLCADYRGVCPMEQAGKRLKTYIRHVF 457
 QY 430 T-LIHQIFDFPQGVVLAHLLEIVQWQMDRSQNPFSQVASYFLQKQKTIQDISHTVYN 488
 DB 438 NSAGTSPVFNKNGDAPGRYDFIYOTNTTPGRYRLIGQWTDLOL-NIEMQMGKGV 516
 QY 489 TTPMSCKRCQSGQKKRKGIVHCCFECIDCLPGFLNHTDEYEQACPNNSYQSE 548
 DB 517 EIPSSVCLPKRPGQKRTKQKTPCCWTCERC--DGVOYQDENTCQCHQPDQRPNEKR 573
 QY 549 TSCFRQLVLEHHEAPTLAVALAALGFLSTLALIVFWRHFQRPPIYRSAGCPCEML 608
 DB 574 TCCQNIPIIKLHSPWAIPIVFLMLGLIATIFWAFPIRNDPIYASGREISYVL 633
 QY 609 TLLVAWVVPVYQGPVKPCQLGQALPFLCTGICISIAV-----RSFOIVCAFKMS 663
 DB 634 TGIFLCYITFLIMAKPDVAVCSFRVFLGL--GMCISIALLTITNRYIRLFEQKKS 691
 QY 664 REPRAVSVVRYQGPVSNAFITV-LKAVIVVIGMLARPQSHRPDDPKITI-----V 717
 DB 692 TAPRLISPSQLA---ITSLSVQLLGVMFGV-----DEPNIITIDYDEK 736
 QY 718 SCNPYRNSLFTNTSLD-LLSVVGFS-----PAYMKELPTNYEAKFTLSMTE 767
 DB 737 TNPPEQARGVLKCDITDLOILCSLGYSLIMVTCVYAIKTRGVDPENFEAKPIGFTM-- 794
 QY 768 YFTSSVSLCTFEMSAYSGVLTVLDL-----LVTVLNILA-TSIG-YEGPKCYMLTFPE 819
 DB 795 -YTTCTVLAFTPIPFPGTQSAEKLYIQTYTLTISMNLSASVALGMLYMPKYYIIIFHE 853
 QY 820 RNT---PAYFNSMIQGYTM 835
 DB 854 LNVQKRKRSEKAVVYATATM 872
 RESULT 6
 MGR4_RAT
 ID MGR4_RAT STANDARD: PRT; 912 AA.
 AC P31423;
 DT 01-JUL-1993 (Rel. 26, last sequence update)
 DT 15-JUN-2002 (Rel. 41, last annotation update)
 DE Metabotropic glutamate receptor 4 precursor (mg1ur4).
 GN GRM4 OR GPRC4 OR MG1UR4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE=92110002; PubMed=1309649;
 RA Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
 RT "A family of metabotropic glutamate receptors.";
 RL Neuron 8:169-179(1992).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE=93332699; PubMed=833867;
 RA O'Hara P.J., Sheppard P.O., Thøgersen H., Venezia D., Haldeman B.A.,
 McGreene V., Houamed K.M., Thomsen C., Gilbert T.L., Mulvihill E.R.;
 RT "The 11and-binding domain in metabotropic glutamate receptors is
 related to bacterial periplasmic binding proteins.";
 RL Neuron 11:41-52(1993).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: IS WIDELY DISTRIBUTED IN THE CNS. PREDOMINANT
 EXPRESSION IS SEEN IN THE GRANULE CELLS OF THE CEREBELLUM.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLR6.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb.ch/announce/>
 or send an email to license@isb.slb.ch).
 CC EMBL; M92077; -; NOT_ANNOTATED_CDS.
 DR EMBL; M90518; AAA93190.1; -;
 DR PIR; JH0563; JH0563.
 DR InterPro; IPR001828; ANF_receptor.
 DR InterPro; IPR000337; GPCR_Mgt.
 DR Pfam; PF00003; 7tm_3; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PR00248; GPCRMR.
 DR PROSITE; PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE; PS00259; G_PROTEIN_RECP_F3_4; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Multigene family.
 FT SIGNAL 1 32
 FT CHAIN 33 912
 FT DOMAIN 33 587
 FT TRANSMEM 588 610
 FT TRANSMEM 611 624
 FT TRANSMEM 625 645
 FT TRANSMEM 646 656
 FT TRANSMEM 657 675
 FT TRANSMEM 676 699
 FT TRANSMEM 700 720
 FT TRANSMEM 721 750
 FT TRANSMEM 751 772
 FT TRANSMEM 773 785
 FT TRANSMEM 786 808
 FT TRANSMEM 809 821
 FT TRANSMEM 822 847
 FT DOMAIN 848 912
 FT CARBOHYD 98 98
 FT CARBOHYD 301 301
 FT CARBOHYD 454 454
 FT CARBOHYD 484 484
 FT CARBOHYD 569 569
 FT CONFLICT 124 124
 SQ SEQUENCE 912 AA; 101818 MW; 336430EF19B4B57 CRC64;

Query Match
 Best Local Similarity

16.2%; Score 721.5; DB 1; Length 912;
 26.5%; Pred. No. 1e-45;

Matches 240; Conservative 161; Mismatches 349; Indels 155; Gaps 37;
 QY 8 ICSLEFLL--WV--LAEPAN---SDFYLPDGLGLSLANMGMVHNLFLQYPMC 59
 Db 16 LCLLSLXARWVBSLQKPGHMHNSIRIDGDTLGLLPVHGRGS-----EKKAC 67
 QY 60 KEVEKVIQYNLMQAMFAVEEINDDSLPGVLLGYEIVDVCYISNNVQVYFLAHED 119
 Db 68 GELK-KEKGIRLEAMFALDRINDDPLNITIGARIDTC--SRDHALRQSLTFVQ 124
 QY 120 NLPIQDYSNY-----ISRVAVIGPDNSQMYANFLSLLLQIQTSAI 167
 Db 125 AL--TEKDGTEVRKSGSGPPIITKPERVGVIGASGSSVIMANILRKIPQISTAST 182
 QY 168 SDELDRKRPALLRTTPASDHHEAVNQLMHPNMNIIIVLSDYTG-----216
 Db 183 APDLSDSRDFDSRVVPSDTQAQANVDYRALKNMYSTLASSEGSYGVEAFQKS 242
 QY 217 RDNGQLGERVARRDICIAFOETPLTQPNOMNTSEERQLVTVDKQSTARVVVFS 276
 Db 243 RENG-----GVCIAQSVKIPR-EPYTGEDKIKRL-----LETSMAGIIFA 285
 QY 277 PDLTYHFENEVLKQNTG-AVMTASENAID--PVHNLTELGHTFLGIT--QSPV 331
 Db 286 NEDDIRVLEAARRANOTGFHFMGSDSGSKAPVL-RLEEVAE---GAVTLPRMS 340
 QY 332 IPGFSEFRENQPOAGPPLSRT-----SQSYCNQECNCLNATLSFYTLRLSG 381
 Db 341 VRGDRDFFS-----SRITDNNRNIMPFEWEDNTHCKLSRALKSGIKKCTN 390
 QY 382 -ERV-----VYSYSAVYAAVAHLSLGCDSKSTCTRY-----VYPMQL 421
 Db 391 RERIGQDSAYEORGVQFVDVAVAMGHALHAM---HRDLPGVAVGLCPMDPDVDTQL 447
 QY 422 EEIWKVNT-LLDHQIFPDQVALLLEIVOMQMSQNPQSVASYP-LQROLKNIQ 479
 Db 448 KYIRNVNPSGIAGNPVYFNENGDAFGYDIYQYOLRNGSAEYKIVGWTHTLHRIEMQ 507
 QY 480 DISWTVNNTIPMSCKRQSGQSKKPKVGIHVCCFECIDCLPDTFLNHTDEYECACP 539
 Db 508 ---WPGSQQLPSPISGSLPCQPGERKKTVMGCMCCERC---TGVOYQVDRTYKTCIP 561
 QY 540 NNEKSYSESCRRQLVLEMEHAPITVAALLAAGFLSTLALVYFMHFPQPIYRSA 599
 Db 562 YDKMPETNRTSCQPIPIVKKLEMDSPMAVLDPLFAVVGIAATLVVYVFNNDPIYKAS 621
 QY 600 GPMCFICLLILLVAVVWVYVYVGPVKVSTLCRQALFPCTFTICSCIAVRSFOICAF 659
 Db 622 GRELSYVLGIFLCVATFTLMAEPDLGCSLRIRPILGMSSTYALLTKNRIRYIF 681
 QY 660 KMASRFPRAVSYVVRQGP--VYSMAETV-LKMYIVIGMLARPOSHPRTPDDPKRTT 716
 Db 682 EOKGRSVSA-----PRTSPASOLAITFILSLDGLGCVFVVDP--SHSVVDQDGR-- 733
 QY 717 VSCNPNRNRLNLTSLD-LLSVGRS-----FAFMGKELPNVNEARFYTISMT 766
 Db 734 -TLDPRAFGVLCDSIDSLICLGLYSMLMVTCTVYAKRTGVPTEFEARPIGFTM- 791
 QY 767 FYFTSSVSLCTFMSAYSQ-----VLVTIVDLLVYVNLIAISIG-YFPPKCYMILFYP 818
 Db 792 --YTTCTVLAFLPIFGTISQSDAKLYITTTLVSVLSASVSLGMLYAPKYYIIILFHP 849
 QY 819 ERNTP 823
 Db 850 EQNVP 854
 RESULT 7
 MGR7_HUMAN
 ID MGR7_HUMAN STANDARD; PRT; 915 AA.
 AC 014831;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

15-JUN-2002 (Rel. 41, last annotation update)
 DE Metabotropic glutamate receptor 7 precursor (mglur7).
 GN GRM7 OR GPRC1G OR MGLUR7.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96437220; PubMed=8840028;
 RA Makoff A., Pilling C., Harrington K., Emson P.;
 RT "Human metabotropic glutamate receptor type 7: molecular cloning and
 RL brain distribution in the CNS";
 RL Brain Res. Mol. Brain Res. 40:165-170(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96141892; PubMed=9473604;
 RA Wu S., Wright R.A., Rockey P.K., Burgett S.G., Arnold J.S.,
 Rostock P.R., Jr., Johnson B.G., Schoepf D.D., Belagaje R.M.;
 RT "Group III human metabotropic glutamate receptors 4, 7 and 8:
 RT molecular cloning, functional expression, and comparison of
 RL pharmacological properties in RGT cells";
 RL Brain Res. Mol. Brain Res. 53:88-97(1998).
 RN [3]
 RP VARIANT PHE-433.
 RX MEDLINE=21095249; PubMed=11163549;
 RA Boltona A.A., Kerwin R.W., Munro J., Arranz M.J., Makoff A.J.;
 RT "Polymorphisms in the genes for mglur types 7 and 8: association
 RL studies with schizophrenia";
 RL Schizophren. Res. 47:99-103(2001).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 CC ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY AREAS OF THE BRAIN,
 CC ESPECIALLY IN THE CEREBRAL CORTEX, HIPPOCAMPUS, AND CEREBELLUM.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR4.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X94552; CA64245.1; -;
 DR EMBL: U92458; AAB51763.1; -;
 DR Gene: HGNC:4599; GRM7.
 DR MIM: 604101; -;
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_REC_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_REC_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_REC_F3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_REC_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Multigene family; Olfaction; Polymorphism.
 FT SIGNAL 1 32
 FT CHAIN 1 915
 FT DOMAIN 33 590 METABOTROPIC GLUTAMATE RECEPTOR 7.
 FT TRANSSEM 591 615 I (POTENTIAL).
 FT DOMAIN 616 627 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 628 648 II (POTENTIAL).
 FT DOMAIN 649 654 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 655 675 III (POTENTIAL).
 FT DOMAIN 676 702 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 703 723 IV (POTENTIAL).

FT DOMAIN 724 753 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 754 775 V (POTENTIAL).
 FT DOMAIN 776 788 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 789 810 VI (POTENTIAL).
 FT DOMAIN 811 825 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 826 850 VII (POTENTIAL).
 FT DOMAIN 851 915 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 486 486 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 572 572 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 433 433 Y->F.
 FT SEQUENCE 915 AA; 102250 MW; CFF94E0BFF7F4919 CRC64;
 SO
 Query Match 16.2%; Score 721.5; DB 1; Length 915;
 Best Local Similarity 27.9%; Pred. No. 1e-45;
 Matches 256; Conservative 152; Mismatches 354; Indels 157; Gaps 41;
 QY 9 CSLPFLVLAEPANSPFLP-----GDYLLGGLSLHANMKIVHLNFIQVCKEY 62
 DB 19 CVLEVLICALAAAGQETAPHSIRIGDYLGLGFLPVAH--KGP-----SGVF-CGDI 70
 QY 63 EVKVIQYMLQMRRAVEINNDSSILPGVLYGEIVQVYISNNVQVLYLAHEDNLL 122
 DB 71 K-RENGIRHLEMLYALQINSDPMLRPVLTIGANILDTG--SRDTYALQSLTFVQAL- 126
 QY 123 PLEDYSNY-----ISRVAVIGPDNSVMTVANFLSLFLLPDTYSASIDE 170
 DB 127 -IQKTSVRCNGPBPVFKPEKVVGYAGSGSVSIWVANILFLFQIPQISTASTAPE 185
 QY 171 LDKRFPALLRTPTSDAHNEAMQMLHPWMNIIYVLSDTYGRNGQLGSRVARR 230
 DB 186 LSDRRYDFFSKVPDPDFQADAMVDYKALGMWNVSTLASGSYG-EKGVESFTQISKE 244
 QY 231 --DIIAFOETLPTLPQNMNTSEEROLY---IYVDKLOQSTARVVVSPDLTLYHF 284
 DB 245 AGGLCIASVRLP-----QERKDTIDPDRITKLOLTPNRAVAVIFANDDINOI 295
 QY 285 FNEVLRNFTG-AVWIASBSNA--IDPVHLNTELGHLGTFGLTIO--SVIPGFSERR 339
 DB 296 LAARARAOVGHFTLWVGSWSKINP-LHQHEDIAE---GAIIQPRARVEBFDAYF 350
 QY 340 EMGPAGPPLSRSTOSYTCN-----QEDCNC--LNATLSFTILRLSG-ERV---- 384
 DB 351 -----TSRTLENNRRNWRFAEYWEENFNCULTISGSKKEPDKCTGQERIGDS 400
 QY 385 -----YYSVYSAVYAVALHSLGCDKSTCTK-RVYYP-----WQLEELMKYNF 429
 DB 401 NTEQSGKVQFVIDAYAAAHALHNM--NKDLCAHYRGVCPREMGAGSKILKTYIRNVF 457
 QY 430 T-LLDHQIFPDQGVVALHLEIVQWQMRSONPFQSVASYPLQRLQNIODISHTVNN 488
 DB 458 NSAGTAPWFMFNKNDAPGRYDFIQYTNTSNPGRYLIGQWDEQL-NIEMQMGKQVR 516
 QY 489 TTPMSCKRSQSGOKKKRPVGIHVCCFPCICDLPFTPLNHTPDEVCQACPNNEKSYSE 548
 DB 517 ELPASVCTLPCKPGQKKTKQGTGPCWTCCEP---DGTYQYQDENTCQHCPTDQRPNEWR 573
 QY 549 TSCFRAQVLEHMEAPTIAVALLAALGFLSLALIVLFMRFQPIYRSAGPACFTML 608
 DB 574 TCCQDIPILKLEHSPMAVIVPFLAMGLIINFMAFIRKRNDFIYRASRELSYVL 633
 QY 609 TLLVAWVAVPYVGPVRYSTCLQALPFLCTICISCIAY-----RSQIVCAFKNAS 663
 DB 634 TGIFLCYITTEFLMKPVAACSFRRVFLGT--GMCISAAALFTKRNIRYRFEQGRKSV 691
 QY 664 RPRAYSWVRQSGYVEMAFITV-LKMVIVIGMLARQSHRRPRDPDKTTI-----V 717
 DB 692 TAPRLISPTSLA--TTSLSVQLLVGFIMFGV-----DPPNIIIDYDEHK 736
 QY 718 SCNPYRSLTFTSLD-LLSVVGFS-----FAYMGKELPTYNNAEFTILSMWF 767

Db 737 TNPBQARGVAKCDITDLOITCSLGLMTCTVYAIKRGVFNENAKPIGFTM-- 794
 QY 768 YFTSSVSLCTFMASGVLTIVDL-----LTVLNLIA-LSLG-YFGPKCYMLFYPE 819
 Db 795 -YTTCTIWLAFIPFIPGTQASAKIYQITTLISNLSASVALGMYKMYIIIFHPE 853
 QY 820 RNF---PAYFNSMIQGYTM 835
 Db 854 LNVQKRKRSEKAVVATATM 872

RESULT 8
 MGR3_RAT
 ID MGR3_RAT STANDARD: PRT: 879 AA.
 AC P31422;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Metabotropic glutamate receptor 3 precursor (mglur3).
 GN GNM3 OR GPRC1C OR MGLUR3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxId:10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA MEDLINE=92110002; PubMed=1309649;
 RX Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
 RT "A family of metabotropic glutamate receptors."
 RL Neuron 8:169-179(1992).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLYLATE CYCLASE
 ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: IS WIDELY DISTRIBUTED IN THE CNS. PREDOMINANT
 EXPRESSION IS SEEN IN THE NEURONAL CELLS OF THE CEREBRAL CORTEX,
 DENTATE GYRUS, AND GLIAL CELLS THROUGHOUT BRAIN REGIONS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR2.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M92076; -; NOT_ANNOTATED_CDS.
 DR PIR: JH0562; JH0562.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PRO0248; GPCR_MGR.
 DR PROSITE: PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_RECP_F3_4; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KM Multigene family.
 FT SIGNAL 1 22
 FT CHAIN 23 879
 FT DOMAIN 23 879 METABOTROPIC GLUTAMATE RECEPTOR 3.
 FT TRANSMEM 576 576
 FT DOMAIN 577 599 I (POTENTIAL).
 FT TRANSMEM 600 613 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 614 634 II (POTENTIAL).
 FT DOMAIN 635 645 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 646 664 III (POTENTIAL).
 FT TRANSMEM 665 688 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 689 709 IV (POTENTIAL).
 FT DOMAIN 710 734 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 735 756 V (POTENTIAL).
 FT DOMAIN 757 769 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 770 792 VI (POTENTIAL).
 FT DOMAIN 793 802 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 803 828 VII (POTENTIAL).
 FT DOMAIN 829 879 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 879 AA; 98959 MW; 3E5965DE56DEED CRC64;

Query Match 16.2%; Score 718; DB 1; Length 879;
 Best Local Similarity 26.1%; Pred. No. 1.8e-45;
 Matches 225; Conservative 147; Mismatches 357; Indels 134; Gaps 29;

QY 29 LPPDYLIGLGFSLHANMKGIVHLNLOVPCKREYKVGISYNNLMQMRFAVEELINDSSL 88
 Db 35 IEGDLVYGGFLPINEKGTGEECGRIEDR-----GIQRLNMLPAIDELINDNL 85
 QY 89 LPGVLLGEYEIVDC-----YISNNVQVLYFLAHEDN---LDP-----IOEDYSNTISRV 135
 Db 86 LPGVKLVGHILDTCSRDYVALLEQSLFEVRASLTKVDEAEVCPDGSVAIGE---NIPILI 142
 QY 136 VAVIYGPNSSEWTVANFLSLFLLPQITTSASIDELBDKRFALLRTPTPSADHRYEAMV 195
 Db 143 AGVIGGSYSVSIVANLRLFPQISVASTAKLSDKSRDYFAFATVPDPFYQAKAMA 202
 QY 196 QLMHFRMNITVLYVSDTFRGNGOLGGERVARDICIAFOETLPIQNONMTSPERO 255
 Db 203 EILRFNMTVTVSVASGSDYGEIGIEAFEDBALRNICITAEKVGSRNIRKSYDS---- 258
 QY 256 RLVTYIDKIQOSTARVAVVE---SPDLTLHFENEVLQNFQAVWJASEMAIDPYLHN 312
 Db 259 ---VIRELQKPARAVYVLEMRSDSRELIYAANRV-NASFT---VWASDGMQAQESIVK 311
 QY 313 LTELGHGTGLGITIOSVP-----PEFSFEREPGAGRPPLSRIS 354
 Db 312 GSEHVAVGA-ITLELASHPRQDRYFQSLNPNYNNRNPWFQDFE----- 356
 QY 355 QSYTCN-----QECNCLNATLSEFTILRLSGERVVSVYSAYVAVALHSLG-- 404
 Db 357 QKQCSIQSNKRNRROYCDRLAIDS-----NYEQEKINFYNAVAYAMAHALHKQRTL 411
 QY 405 CDKST--CTRRVYPMQLL--EETWKNVTL-----LDHQTFFDPQGVALLHLEIVQW 453
 Db 412 CPNTTKLCKAMKILIDGKRLKEYLAKINFAPPNPKGADSIYKFDFFGDMGRYVFNFL 471
 QY 454 QMRSONPPOSVASVYPLQRLQKNIQDISMHTVNNITPMSCSKROOSGQKKRPVGIHW 513
 Db 472 QQTGKYSYLKVGHW--AETLSLDVDSIHWS--RNSVPYSQSDPCAPNEMKNNQPDVC 527
 QY 514 CFCICIDLPSTFLNHEDEYEQACPNNEWSYQSESCFROLVFLFEMHEAPTAVALLA 573
 Db 528 CWICICEPEPEYEL--VDEFTCMDGCGQMPFADLSGCYNLIPDYIKWEDAMAIQPVITA 584
 QY 574 ALGFLSLALIVLFWRFQPIYRSAGCPMCLMTLLIYAVVAVPYVVPKPVSTCLR 633
 Db 585 CLGFLCICIVITFEIKNNNPVLKVASGRELCYILLGVSLSYCMTEFFIKPSPVICALR 644
 QY 634 QALFPLCFTCICICINAVSQIYCAF---KMARFPRAYSVWRYQOP---VYSMAFIT 686
 Db 645 RLGLGTSFAICYSALLTKTKICARIFDGVANGARREFIS-----PSSQVITCIGLLI 697
 QY 687 VLKNVYVIGMLARPOSHPTDPDKITIVSCNPNTNLSLNTSLDLLSVGSEFAY 746
 Db 698 VQIVAVSWVILFETPGRRATLPEKRETVLKCNVK--DSMLSLSTFVAVLLICVYAF 756
 QY 747 MGEKLPINVEAKFTISMTFTSSVSLCTFMSASGVLTIVDLYL-IVLMLAISLG 805
 Db 757 KTRKCEPENEAKPIGFTM--YTTCTIWLAFIPF---YFTSSDYRVQTTMCISSVLS 810
 QY 806 -----YFGPKCYMLFYPERN 821

Db 811 GFVVLGCLFAPKRVHIVLFOQRN 833

RESULT 9

MGRA_HUMAN STANDARD; PRT; 912 AA.

AC Q14833;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Metabotropic glutamate receptor 4 precursor (mgIur4).

GN GRM4 OR GPRC1D OR MGLUR4.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

OX (1)

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=96346635; PubMed=8738157;

RA Makoff A., Leichuk R., Oker M., Harrington K., Emson P.;

RT "Molecular characterization and localization of human metabotropic glutamate receptor type 4.";

RL Brain Res. Mol. Brain Res. 37:239-248(1996).

RL (2)

RN SEQUENCE FROM N.A.

RX MEDLINE=98141892; PubMed=9473604;

RA Wu S., Wright R.A., Rockey P.K., Burgett S.G., Arnold J.S., Roseck P.R. Jr., Johnson B.G., Schoepf D.D., Belagaje R.M.;

RT "Group III human metabotropic glutamate receptors 4, 7 and 8: molecular cloning, functional expression, and comparison of pharmacological properties in RGT cells.";

RL Brain Res. Mol. Brain Res. 53:88-97(1998).

RL (3)

RN SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=95342351; PubMed=7617140;

RA Flor P.J., Lujck S., Rueegg D., Leonhardt T., Knoepfel T., Kuhn R.;

RT "Molecular cloning, functional expression and pharmacological characterization of the human metabotropic glutamate receptor type 4.";

RL Neuropharmacology 34:149-155(1995).

RL (4)

RN VARIANT ILE-797.

RP MEDLINE=21416233; PubMed=11525421;

RA Ohtsuki T., Toru M., Arihama T.;

RT "Mutation screening of the metabotropic glutamate receptor mglur4 (GRM4) gene in patients with schizophrenia.";

RL Psychiatr. Genet. 11:79-83(2001).

CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLYLATE CYCLASE ACTIVITY.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN THE CEREBELLUM. EXPRESSED AT LOW LEVELS IN HIPPOCAMPUS, HYPOTHALAMUS AND THALAMUS. NO EXPRESSION DETECTED IN LIVER.

CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS. STRONGEST, TO MGLUR6.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

CC EMBL: X80818; CAAS6784.1; -

DR EMBL: U92457; AAB51762.1; -

DR Genew: HGNC:4596; GRM4.

DR MIM: 604100; -

DR InterPro: IPR001828; ANF_receptor.

DR InterPro: IPR000337; GPCR_Mgr.

DR Pfam: PR00003; 7tm_3; 1.

DR Pfam: PR00248; ANF_receptor; 1.

DR PRINTS: PR00979; G_PROTEIN_REC_P3_1; 1.

DR PROSITE: PS00979; G_PROTEIN_REC_P3_1; 1.

DR PROSITE: PS00980; G_PROTEIN_REC_P3_2; 1.

DR PROSITE: PS00981; G_PROTEIN_REC_P3_3; 1.

DR PROSITE: PS00982; G_PROTEIN_REC_P3_4; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;

KW Multigene family; Polymorphism.

FT SIGNAL 1 32

FT CHAIN 33 912

FT DOMAIN 33 587

FT TRANSMEM 588 610

FT DOMAIN 611 624

FT TRANSMEM 625 645

FT DOMAIN 646 656

FT TRANSMEM 657 675

FT DOMAIN 676 699

FT TRANSMEM 700 720

FT DOMAIN 721 750

FT TRANSMEM 751 772

FT DOMAIN 773 785

FT TRANSMEM 786 808

FT DOMAIN 809 821

FT TRANSMEM 822 847

FT DOMAIN 848 912

FT CARBOHYD 98 98

FT CARBOHYD 301 301

FT CARBOHYD 454 454

FT CARBOHYD 484 484

FT CARBOHYD 569 569

FT VARIANT 797 797

FT V -> I.

FT VARIAT -VAR_012992.

FT FTID=36E63A2EAF5A CRC64;

SQ SEQUENCE 912 AA; 101867 MW; 4A2F36E63A2EAF5A CRC64;

Query Match 16.1%; Score 717.5; DB 1; Length 912;

Best local similarity 26.1%; Pred. No. 2, 1e-45;

Matches 224; Conservative 164; Mismatch 362; Indels 135; Gaps 35;

QY 8 ICSLFFLL--WV--LAEPAN--SDFYLPQDYLLGGFLSLHANKGIVHLNFIQVPM 59

DB 16 ICSLFLSGPWPSSSGKPKGHPHMSIRIDGITGLGEPVHGKRS-----EGKPC 67

QY 60 KEVEAVVIGNLMQARFAVEELINNOSSILPGLLGEIYDVYISNNQPVLYFLAHED 119

DB 68 GELK-KEKGIHRLAMFLALDRINDPDLPTTIGARILDTG--SRDTHALQSLTFVQ 124

QY 120 NLPIQEDYSNY-----ISRVAVIGPDNSESVYVAVNLSFLPQIYTSAT 167

DB 125 AL--IKKDGTEVAGCGSGPPIITKPERVGVIGIASSSVSIWVNIIRLFKTIQISYAS 182

QY 168 SDELDRKVRPALLRTTPSADHHEAMVOLMIFRNWIIIVSSDPTGDCNQLGGERV 227

DB 183 APLDSNSRDEFSRVPSPDYQAQAVDYLRAKNVYSTVASSESGSGVEAFIQR 241

QY 228 ARRD--ICIAFQETLTLPQNMNMTEERQRLVITIDKLOQASRVRVYSPDLTLXHF 285

DB 242 SRDGGVCTAQSVPKPR-EKAGEFDKIIRL-----LETNARAVITIPANDDITRVL 294

QY 286 NEVLRONFTG-AVWIASESWA--IDPVLNLTDELGHGIFLGTI--QSVPIPFSEFPE 340

DB 295 EAARRANQGHFPMWSDSGSKIAVYL-LEVAE-----GAVTILPKRSVAGPFRYS 349

QY 341 WGPQACPPPLSRT-----SQSYTCNQDCNCLNLTSLNTLRLSG-ERV----- 384

DB 350 -----SRLDNNRRNIMFAEFWEDNFHCKLSRHALKKGSHVKKCTNRRIQODSA 399

QY 385 -----VYSYSAVVAVAHSLGCDKSTCKRY-----VYPMOLLEIMKVNFT 430

DB 400 YEOEGKVQVFIADVAVYAMGHALHAM--HDLQCGRGVGLCPRRDPVGTOLKTIKRVNS 456

QY 431 -LIDHOLFDPDQGVALLHLEIVQWMDRSONPQSVASYYPILQRLKNIQDISMHTVNT 489

Db 457 GIAGNPVTENENGAPGRYDIYQYOLNRNDSAEYVIGSW--TDHLHLRIEMHMPGSGQO 514
 QY 490 IIPSMCKROSGOKKPVGIHCVCEFCIDLPGETFLNHDEDEYCOACPNMNSYSOSET 549
 Db 515 LPSISICSLPQOPGRKRTVGMPCMCNCEPC---TGYYQVDRYTCCTCPDMRPENTRT 571
 QY 550 SCRRQLVFLFEMHAPPIAVALAALGLSTLALVLFEMHFPQPIYVSAGGPMCLMLT 609
 Db 572 GCRPIIKLEMGSPMAVLPLFLAVGIALMTLVITFVRXNDPIYKASGRELSTYLLA 631
 QY 610 LLVAVVAVVYVPPRVYVCLCRQALFPLCFITICICIAVRSQYICAKMASRFRAY 669
 Db 632 GFLCYATFELMAEPDLGTCISLRIFLGMSISYALLTNRIRYIRIFEQGRSYA- 690
 QY 670 SYWRVYQGPVYMAF---ITVLKAVIVIGLARPOSHPRTPDPDPKITIYSCPNRYS 726
 Db 691 ---PRFISPASQALITSLISLQIGCVFVDP-SHVVYEDQDQ---TIDPRARG 742
 QY 727 LLENTSLDL-LLSVGFS-----FAYMKELPTNVEAKPITLSMTFTYSVSLC 776
 Db 743 VLKCDISDLISLILGLSMLMTVCTVYAKTRGVPETFEAKPIGFTM---YTTCIYWL 799
 QY 777 TEMSAYSG-----VAVTIVDLVYVNLNLALISLG-YRQPKCMITFPERNTP 823
 Db 800 AFPIPIFGTSQASADKIXIQTTLTVSVLSASVSLGMLMPKYYIILFHEQNPV 854

RESULT 10

MGR3_HUMAN STANDARD; PRT; 877 AA.

AC Q14832;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 3 precursor (mglur3).
 GN GPM3 OR GPRC1C OR MGLUR3.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96437205; PubMed=8840013;
 RA Makoff A., Volpe F., Leichuk R., Harrington K., Emson P.;
 RT "Molecular characterization and localization of human metabotropic
 glutamate receptor type 3.";
 RL Brain Res. Mol. Brain Res. 40:55-63(1996).
 CC -FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLYLATE CYCLASE
 ACTIVITY.
 CC -SUBCELLULAR LOCATION: Integral membrane protein.
 CC -SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: X77748; CAA54796.1; -
 DR Genew; HGNC:4595; GRM3.
 DR MIM; 601115;
 DR InterPro; IPR001828; ANF_receptor.
 DR InterPro; IPR000337; GPCR_Mgr.
 DR Pfam; PF00003; 7tm_3; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PR00248; GPCRMR.
 DR PROSITE; PS00979; G_PROTEIN_RECP_F3_1; 1.

DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE; PS00259; G_PROTEIN_RECP_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KM Multigene family.
 FT SIGNAL 1 20
 FT CHAIN 21 877
 FT DOMAIN 21 574
 FT TRANSMEM 575 597
 FT DOMAIN 598 611
 FT TRANSMEM 612 632
 FT TRANSMEM 633 643
 FT TRANSMEM 644 662
 FT TRANSMEM 663 686
 FT TRANSMEM 687 707
 FT TRANSMEM 708 732
 FT TRANSMEM 733 754
 FT TRANSMEM 755 767
 FT TRANSMEM 768 790
 FT TRANSMEM 791 800
 FT TRANSMEM 801 826
 FT TRANSMEM 827 877
 FT CARBOHYD 207 207
 FT CARBOHYD 290 290
 FT CARBOHYD 412 412
 FT CARBOHYD 437 437
 SQ SEQUENCE 877 AA; 9619 MW; 66F28663CE35F740 CRC64;

Query Match

Score 709; DB 1; Length 877;
 Best Local Similarity 25.8%; Pred. No. 8.4e-45;

Matches 224; Conservative 145; Mismatches 356; Indels 142; Gaps 28;

QY 29 LPQDYLGLGFLSHANKKGIIVHLNFILOVPMCKEYEVVIGYNLMQANFAVEIINDSSL 88
 Db 33 IEGDLVYLGFLPIINEKGTGTEECGRINEDR-----GIQRLAMLAIDEINKDYL 83
 QY 89 LPQVLYGEIYDVC-----YISNNVQVLYFLAHEDN---LLP-----IQEDSNYSIRV 135
 Db 84 LPQKLVHILDLTCSRDYALLESLEFVRASLTKRVDAEYMCDPGSAIIDE---NIPLLI 140
 QY 136 VAVIGPNSSESVMTVANFLSLFLPQITYSAISDELDRKYRFPALNTPTTSADHVEAMV 195
 Db 141 AGVIGGSYSVSIVQANLNLFLQIPQISYASTAKLSKDRYDFEATVPDPFQAQAMA 200
 QY 196 QLMHFRNMNIIYVSSDYGRDNGQLIGERVARDCICAPQETLPTLPQNPQNTSEKQ 255
 Db 201 EILTFEWMYTVSVASGDEGEIEAFQEARLRNICITAEKRVGSNIRKSYDS---- 256
 QY 256 RLVTIYDKIOSTARVYVSPDLTLHFNEVILRON--FTGAVWISSEMAIDPVILNL 313
 Db 257 ---VIRELQKPNAVVVLFMRSDSRELIAAASRAVASFT---VWASIDMGAOESILKG 310
 QY 314 TELGLHGFPIGITVQSVPI-----PGSEPREMPQAGPPLPSTSQ 355
 Db 311 SEHVAIGA-ITTELASQPVQDFRYQSILNPNHNPFRDWE-----Q 355
 QY 356 SYTCNDE-----CNCINATLSENTIILRSERYVSVYSVAVAHALSHLSLG--C 405
 Db 356 KFGCSLQNKNNHRYCCKHALIDSS-----NYDESKIMEVNAVYAMAHALHKMQTTLIC 410
 QY 406 DKST--CTKRVVYVWOLL--EELMKVNTL-----LDHQIFPDQGVALLHLEIYQWQ 454
 Db 411 PNTKLCDARKILDKGLKYLDYLLKINFYLPFNPNKADSIYKFDGDMGRNVFN-- 468
 QY 455 WDSQNPFGSVASYLPQR-----QLNKIDISNHTYNNNTIPMSCKROSGOKKRPVG 509
 Db 469 -----FQNVGKYSTLKGHMAETLSDVNSTIHSRNSVPTSQCDPCAPNEMKMOP 521
 QY 510 IHVCCFECIDLPGETFLNHDEDEYCOACPNMNSYSOSETSCRRQLVFLFEMHAPPIAV 569
 Db 522 GDVCCWICIPCPEYEL---ADEFTCMDCSGQMPFADLTGCTDLPEDYIRWEDAMALIGP 578
 QY 570 ALTAALGLSTLALVLFEMHFPQPIYVSAGGPMCLMLTLVAVVAVVYVPPRVST 629

Db 579 VTIACLGMCCTCMVYTVIKHNHNPVLYKASGELCYIIILFGVGLSYCTFFPIAPSPYI 638
 QY 630 CIGCRALPELFTTICISCIAYRSFQIVCAF---KMASFPFAPSYWVRYOGP---YVSM 682
 Db 639 CALRRLGSSGFSALCYLFTKNCIARLPDVKNGAQRPKIS-----PSSGVFICL 691
 QY 683 AFITVLMKAVIYVIGMLARQSHPRDPDKITIVSCNPNTNLSLNTSLDILLSVYGF 742
 Db 692 GLIIVQIVWVSWMILLEAPGRTYRLAKRETVILKCNVK-DSSMLISLIVYIILICT 750
 QY 743 SFAYWKEPLTNYNKAFTLTSMFTYSVSLCTFMSASGVLTIVDLYL-IVLNLIA 801
 Db 751 VYAFETKRCPEFNENAKTIGFTM---YTTCIMAFLEPIF---YTSSDYRQVTTMCIS 804
 QY 802 ISLG-----YEGPKCYMLLEFPERN 821
 Db 805 VSLSGFVVILGCLFAKRVHIIILFOPQKN 831

RESULT 11

MR2_RAT STANDARD; PRT; 872 AA.

AC P31421;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Metabotropic glutamate receptor 2 precursor (mglur2).
 GN GRM2 OR GPRC1B OR MGLUR2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RX NCBI_TaxID=10116;

SEQUENCE FROM N.A.

RC TISSUE=Brain;
 RX MEDLINE=92110002; PubMed=1309649;
 RA Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
 RL Neuron 8:169-179(1992).

-1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
 MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE ACTIVITY.
 MAY MEDULATE SUPPRESSION OF NEUROTRANSMISSION OR MAY BE INVOLVED IN
 SYNAPTOGENESIS OR SYNAPTIC STABILIZATION.

-1- SUBCELLULAR LOCATION: Integral membrane protein.

-1- TISSUE SPECIFICITY: IS WIDELY DISTRIBUTED IN THE CNS AND PROMINENT
 EXPRESSION IS SEEN IN GOLGI CELLS OF THE CEREBELLUM AND SOME

PARTICULAR NEURONAL CELLS IN OTHER BRAIN REGIONS.

-1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 SPRONGEST, TO MGLUR3.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL; M92075; -; NOT_ANNOTATED_CDS.
 DR PIR; JH0561; JH0561.
 DR InterPro; IPR001828; ANF_receptor.
 DR InterPro; IPR000337; GPCR_Mgr.
 DR Pfam; PF00003; 7tm_3; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PR00248; GPCRMR.
 DR PROSITE; PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE; PS0259; G_PROTEIN_RECP_F3_4; 1.
 KM G-protein coupled receptor; transmembrane; Glycoprotein; Signal;
 KW Multigene family.
 FT SIGNAL 1 18 POTENTIAL.

FT CHAIN 19 872 METABOTROPIC GLUTAMATE RECEPTOR 2.
 FT DOMAIN 19 567 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 568 590 I (POTENTIAL).
 FT DOMAIN 591 604 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 605 625 II (POTENTIAL).
 FT DOMAIN 626 636 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 637 655 III (POTENTIAL).
 FT DOMAIN 656 679 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 680 700 IV (POTENTIAL).
 FT DOMAIN 701 725 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 726 747 V (POTENTIAL).
 FT DOMAIN 748 760 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 761 783 VI (POTENTIAL).
 FT DOMAIN 784 793 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 794 819 VII (POTENTIAL).
 FT DOMAIN 820 872 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 547 547 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 872 AA; 95773 MM; 1874CABD6AD4BED9 CRC64;

Query Match 15.7%; Score 699.5; DB 1; Length 872;
 Best Local Similarity 26.8%; Pred. No. 4,2e-44;
 Matches 234; Conservative 157; Mismatches 365; Indels 117; Gaps 33;

QY 14 LLM-VLAEPENSDFYLDGYLLGGLFSLH-----ANNKGIYHFLQVPMCKEYEVYI 67
 Db 12 LLMGVAEGPAKKVYLTLEGDVILGGLFPVHQGGAEECGPVN-----EHR----- 57
 QY 68 GYNIMQAMRFVVEEINNDSILLPGVLLGLEYIDVQYISNN-VQVLYFLAHE-----DNL 121
 Db 58 GIGRLAAMFLADRLINRPHLPGLVRLGAHLITDSCSDKTHALQALDVPVRLSLSGAGDS 117
 QY 122 LPIQED-----YSNTISRVAVITGPDNSESVTVAFLFLPQITYSATSDRLDRVR 176
 Db 118 RHICPDGSYAHNSDAFVAVTGIGSYSDVSIOVANILRLIQIDISYASISAKISDSR 177
 QY 177 FPAALRTPSADHVEAMVQLMHRMWTIYVSSDPYGRDNGOLLGERVARDICIAF 236
 Db 178 YDFARVTPPEFFQAKAAEILRFENMTYVSIVASEGDYGEGLIAFELEAARNICVAT 237
 QY 237 QETLTLPQNMNTEEBORLVTIYDKLOQSAARVYVE--SPDLTFHFENEVROMT 294
 Db 238 SEKV-----GRMSRAAEGVYRAL--LQKPSARAIVLTSSDKRELATLATORLNAST 290
 QY 295 GAVWIASSEW-AIDFVLNLTGLHGTFLGTTIOSVPIGFSEERE---WGPOAGPPP 349
 Db 291 ---WVASDGMGLSESVAG-SERAEGA-ITIELASYPISDPASYFQSLDPMNSRNPMF 345
 QY 350 LSRTSQSTCNOECNCLNATISFTTIRLGERVYVSVAIVAVARALSL--LGCDK 407
 Db 346 REFWEERHCFRQDCAHSL---RAVPFDESKIMVVAVMAHALNHRALCPN 402
 QY 408 ST--C-TRVYVPMOLLEE-IMKVNFYTL-----LDHOIFPQDQVALHLEIYQW-OMD 456
 Db 403 TTHLCDAMRPVNGKRLYDFVLNKFDAFPFPADIDDEYRDRFSDGIGRINIFYYLAG 462
 QY 457 RSONPQSVASYYPQLQROLKNIODISMHTVN-NTIPMSGSKRCGOKRRPVGIHVCF 515
 Db 463 SGRVYQKVG--YMAEGILDTISFIPMASPAGPLPARCSGPCQNVKVSQPEVCW 520
 QY 516 ECIDLPCTFLNHTDEYEQACPNNEYSQSESCFRQVLFLEMHKAPITIAVALLAL 575
 Db 521 LCIPQCP---YEYRDEFTCADCGIGWPNASLTCCFELPEYIRWGMAMVAVPTIACL 577
 QY 576 GFLSLATLVIFWHRFOPPIYRSAGGPMCLMTLLVAAYMVVYVGPVKVSTLCRQA 635
 Db 578 GALATLVLAGVFRHNATPVKASGRELCTIILGGVFLCYCTVFIYIAKPSAVCTTLRL 637
 QY 636 LFPPLFTTICISCIAYRSFQIVCAFEMASRFPFAPSYWVRYOGP---YVSMATVLMKV 691

FT TRANSMEM 617 637 II (POTENTIAL).
 FT DOMAIN 638 648 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 649 667 III (POTENTIAL).
 FT DOMAIN 668 693 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 694 714 IV (POTENTIAL).
 FT DOMAIN 715 737 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 738 759 V (POTENTIAL).
 FT DOMAIN 760 772 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 773 795 VI (POTENTIAL).
 FT DOMAIN 796 801 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 802 827 VII (POTENTIAL).
 FT DOMAIN 828 1212 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 378 378 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 734 734 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 877 908 MISSING (IN ISOFORM 5A).
 SQ SEQUENCE 1212 AA; 132468 MW; A3C7360668106A25 CRC64;

Query Match 15.2%; Score 675.5; DB 1; Length 1212;
 Best Local Similarity 25.7%; Pred. No. 3.9e-42;
 Matches 230; Conservative 153; Mismatches 333; Indels 179; Gaps 34;

QY 28 YLPDYLGLGLFSLHANMKGIVHLNLPVMPCKEYEVK-----IGYNLQAMRFAYEE 81
 DB 30 HMPGDIIIGALFVH-----H-----QPTVDYHERKCGAVRQYGIQVREAMLHTLER 78
 QY 82 INDDSLIPGVLLGYELVDYCYISN-NVQVLYFLAHEDNLPIQED-----YSN 130
 DB 79 INSDPTLPITITGCEIRDCMHSVALEQSIIEFI--RDSLISEEEGLVRCVDSGSSS 136
 QY 131 YISR--VAVIGPDSSEVTVANFLSLFLPOTYSASIDELDKVFPALLRTTPSAD 188
 DB 137 FRSKPIYGVIGGSSVAQVONLQLENIPIQAIYATSDLSKTLFFKPYRVPSPA 196
 QY 189 HVEAMVQMLHFRMNIIVLVSSDYGRDNGQLGERVARDCIAFOETLPIQNPON 248
 DB 197 QQARAWDIKRYNMTVVSIVHTEGNAGESMEAFKMSAKEGICIAHSYKIYSNAGEOS 256
 QY 249 MTSEERRLTYIDKLOOS-----TARYVVPSPDLTIHFNEVLRQNFQA-VVIAS 301
 DB 257 F-----DKLTKLTSHLPKARVYACFCGEMTVRGLMLAMRRLLAGLLELLGS 304
 QY 302 ESWAIDPVNLHTELGHTGLGTTI--QSVPI-----PGSEFEREW 341
 DB 305 DGWA---DRYDVTGQREAVGGITIKQSPDYKMFDDYILKLRPEINHNPMQEFWQH 361
 QY 342 GPQAGPPPLSRSTOSYTCNOECNCLNATLSFNTILRSGERYYS---VYSAVAVAH 397
 DB 362 RFQCRLEGFOENSKY--NKTTC-----NSLTLKTHVODSKMGFVINALYSMAY 409
 QY 398 ALHSILGCKSTC-----TRVYVPMQLLEIKVNFNT-LLDHOFDPQGDVALHL 448
 DB 410 GLHNMM--QMSLPGYAGLDAMKPIDGRKLLSLMKTNFTVSGDTILFDENSDPGRY 466
 QY 449 EIVOMQ-----WDRSONFQSVASYPLQRLKNIDISMTVNTIPMSM 494
 DB 467 EINNFEMGKDYFDYINVSMDNG-----ELKMDDEVMKSKSNIT-RSV 510
 QY 495 CSKRCSGQ---KKRPVGIHVCCFECIDCLPGTFLNHTEDXEYEQACPNNEWSYOSETS 550
 DB 511 CSBPCEKQIKVIRKGEVS---CCWCTCPCKENEYV---FDEYTCRACQQLGSMPTDULTG 564
 QY 551 CFKROLVTEHMEHAPTIAVALLAAGLSLALIVTFMRHPTPIYASAGPACMLMTL 610
 DB 565 CDLIPVOYLKMGPEPIAAVAFACGLGLATLEVTVVFIIIRDPVYKSSSRELCYIITLAG 624
 QY 611 LTVAVMVVYVYVGPVKYSTCLCRQALEPLCTTCISCIASVFOYVCAF-----KMASRF 665
 DB 625 ICGIYICTPCLIAKPKQIYCIQIRIGIGLSPMSYSALVTKTNRIARIILAGSKKICTK 684

QY 666 PRASYVWRYQGYVSMAFITV-LKMWIVVIGMLARPOSHRDPDDPKITIVSCNENYR 724
 DB 685 PRFMSACQ-----LVIAFLICIQIGIYALTIMP-----PD-----IMHDPsir 727
 QY 725 N-SILENTSIDLLISVVGES-----FAYNGKELPIYNNAKFTTSLMTYFPSSVS 774
 DB 728 EYVLICWTNMGVVTPLGYNGLILSCTFYAFRTRNVPANFENAKYIAFTM---YTTCLII 784
 QY 775 LCTFMSASGVLTIVDLYVNLNLASLG-YEGPCYMIIFYPERENTPAYFNS 828
 DB 785 WIAFVPIYFSNKKIITMCFSVLSLSATVALGCMFVRKVIITILAKPERNVSATFT 839

Search completed: May 19, 2003, 09:50:27
 Job time : 15.9089 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:48:15 ; Search time 38.1811 Seconds
(without alignments)
4522.333 Million cell updates/sec

Title: US-09-927-315-9
Perfect score: 4443
Sequence: 1 MGPRAKTICSEFLMWLAEE.....ERNPAYENMIGCYTMRD 838

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs., 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPRMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4392.5	98.9	839	4 O8TE23	O8TE23 homo sapien
2	3203.5	72.1	843	11 O9Z0R7	O9Z0R7 rattus norv
3	3163.5	71.2	843	11 O9Z514	O9Z514 mus musculu
4	3151.5	70.9	843	11 O9Z3J8	O9Z3J8 mus musculu
5	1454	32.7	840	11 O9Z0R8	O9Z0R8 rattus norv
6	1441	32.4	842	11 O9Z515	O9Z515 mus musculu
7	1440	32.4	842	11 O9Z515	O9Z515 mus musculu
8	1434	32.3	842	11 O9Z5J9	O9Z5J9 mus musculu
9	1431	32.2	842	11 O9Z5J9	O9Z5J9 mus musculu
10	1253.5	28.2	763	4 O8TDJ9	O8TDJ9 homo sapien
11	1062.5	23.9	940	13 O73635	O73635 fugu rubrip
12	1039	23.4	858	11 O9Z3K1	O9Z3K1 rattus norv
13	1030	23.2	940	13 O9Z0R6	O9Z0R6 spatius aura
14	1018	22.9	858	11 O9Z5D8	O9Z5D8 mus musculu
15	1014	22.8	858	11 O9Z5A4	O9Z5A4 mus musculu
16	1013	22.8	858	11 O9Z3K0	O9Z3K0 mus musculu

17	1012.5	22.8	877	13 O9PW88	O9PW88 carassius a
18	1012	22.8	858	11 O9Z5D9	O9Z5D9 mus musculu
19	1012	22.8	858	11 O91V44	O91V44 mus musculu
20	980	22.1	864	13 O73637	O73637 fugu rubrip
21	975.5	22.0	880	13 O73639	O73639 fugu rubrip
22	925	20.8	875	13 O73640	O73640 fugu rubrip
23	911	20.5	868	13 O73636	O73636 fugu rubrip
24	907.5	20.4	848	13 O93553	O93553 carassius a
25	898	20.2	856	13 O73638	O73638 fugu rubrip
26	824	18.5	912	11 O70410	O70410 mus musculu
27	822	18.5	844	13 O93552	O93552 carassius a
28	744.5	16.8	779	11 O35269	O35269 rattus norv
29	718	16.2	879	11 O9OXS2	O9OXS2 mus musculu
30	710	16.0	877	4 O8TBH9	O8TBH9 homo sapien
31	693.5	15.6	850	11 O35189	O35189 mus musculu
32	693.5	15.6	983	11 O6Z916	O6Z916 rattus norv
33	671.5	15.1	1156	13 O98UC6	O98UC6 gallus gall
34	671.5	15.1	1188	13 O98UC5	O98UC5 gallus gall
35	671.5	15.1	1242	13 O98UC4	O98UC4 gallus gall
36	665.5	15.0	855	11 O70409	O70409 mus musculu
37	654	14.7	1199	11 O9EPV6	O9EPV6 mus musculu
38	653	14.7	866	11 O35268	O35268 rattus norv
39	649	14.6	1218	13 O902F3	O902F3 oncorhynch
40	644.5	14.5	803	11 O35191	O35191 mus musculu
41	636.5	14.3	977	13 O9PWE1	O9PWE1 ictalurus p
42	632.5	14.2	1267	5 O93564	O93564 caenorhabdi
43	617.5	13.9	457	4 O8TDH8	O8TDH8 homo sapien
44	607.5	13.7	808	11 O35190	O35190 mus musculu
45	566	12.7	852	11 O35192	O35192 mus musculu

ALIGNMENTS

RESULT 1

O8TE23 ID O8TE23 PRELIMINARY; PRT; 839 AA.
AC O8TE23:
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Taste-specific G-protein coupled receptor TIR2.
GN TASTIR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21927605; PubMed=11917125;
RA Li X., Staszewski L., Xu H., Durick K., Zoller M., Adler E.;
RT "Human Receptors for sweet and umami taste."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4692-4696(2002).
DR EMBL; AF458149; AAM12239.1; JOINED.
DR EMBL; AF458149; AAM12239.1; JOINED.
DR EMBL; AF458151; AAM12239.1; JOINED.
DR EMBL; AF458152; AAM12239.1; JOINED.
DR EMBL; AF458153; AAM12239.1; JOINED.
KW Receptor.
SQ SEQUENCE 839 AA; 95072 MW; 8CED6F1503F05FDE CRC64;

Query Match 98.9%; Score 4392.5; DB 4; Length 839;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
OY 1 MGPRAKTICSEFLMWLAEEPAENSDFTLPGDYLGGFLSLHANKGYHLNLFLOVPMCK 60
DB 1 MGPRAKTICSEFLMWLAEEPAENSDFTLPGDYLGGFLSLHANKGYHLNLFLOVPMCK 60
OY 61 EYEVAVIGYNLMQAMRFAVEETINDSSLPGVLLGYLGYVDVCYISNNVQPVLYFLAHEDN 120
DB 61 EYEVAVIGYNLMQAMRFAVEETINDSSLPGVLLGYLGYVDVCYISNNVQPVLYFLAHEDN 120

QY 121 LLEIODEYSNYSRYAVVAVIGPDNSESVMYANFLSLFLPQITYSAISDELDRKVRPAL 180
 DB 121 LLEIODEYSNYSRYAVVAVIGPDNSESVMYANFLSLFLPQITYSAISDELDRKVRPAL 180
 QY 181 LRTTPSADHVEAMVQMLHFRNMWIIYVSSDTYGRDNGOLGERRARDICIAFOETL 240
 DB 181 LRTTPSADHVEAMVQMLHFRNMWIIYVSSDTYGRDNGOLGERRARDICIAFOETL 240
 QY 241 PTLQPNQNMNTSEERQRLVTVYDKLOOSTARVAVVESPDLTYHFENYLRONFGAWIA 300
 DB 241 PTLQPNQNMNTSEERQRLVTVYDKLOOSTARVAVVESPDLTYHFENYLRONFGAWIA 300
 QY 301 SESMAIDPVLANTELHGLTGLTIGTIOVPIPGFSEFEREMGPOAGPPPLSRTSQTGN 360
 DB 301 SESMAIDPVLANTELHGLTGLTIGTIOVPIPGFSEFEREMGPOAGPPPLSRTSQTGN 360
 QY 361 QECNCLNATLSEFTIIRLSGERVYVYSVAVVAHALHSLGCDKSTCKRKYVWOL 420
 DB 361 QECNCLNATLSEFTIIRLSGERVYVYSVAVVAHALHSLGCDKSTCKRKYVWOL 420
 QY 421 LLEIMKVNFTLDDHQIFPDQGVALHLEIYOMQDRSONFQSVASYPLQROLKNTOD 480
 DB 421 LLEIMKVNFTLDDHQIFPDQGVALHLEIYOMQDRSONFQSVASYPLQROLKNTOD 480
 QY 481 ISMHTVNTIPMSCKRSKRGSGQKKKPGVJHVCCECICDLPCTFLNTEDEYECQACPN 540
 DB 481 ISMHTVNTIPMSCKRSKRGSGQKKKPGVJHVCCECICDLPCTFLNTEDEYECQACPN 540
 QY 541 NEMYSQSETCFQRQVLEHHEAPTAVALALGFLSTALIVYWRHQTIVSAG 600
 DB 541 NEMYSQSETCFQRQVLEHHEAPTAVALALGFLSTALIVYWRHQTIVSAG 600
 QY 601 GPMCFMLTLLVAVVAVVYVGPVKYSTCICRQALPPLCTICISCIARVFOIVCAF 660
 DB 601 GPMCFMLTLLVAVVAVVYVGPVKYSTCICRQALPPLCTICISCIARVFOIVCAF 660
 QY 661 MASFPRAYSYVWRYOGPYVSMATYVVKWIVYIGMLARQOS--HPTDPDPKITYSC 719
 DB 661 MASFPRAYSYVWRYOGPYVSMATYVVKWIVYIGMLARQOS--HPTDPDPKITYSC 719
 QY 720 NPNRNSLLEFNTSLDLLSVGFSFAVWKGELPTNYNEAKTITLSMPEYFSSVSLCFEM 779
 DB 720 NPNRNSLLEFNTSLDLLSVGFSFAVWKGELPTNYNEAKTITLSMPEYFSSVSLCFEM 779
 QY 781 SAYSGVLVTVYDLVTVLNLALISLGYFGPKCYMILFPERNTPAYFNSMIGYTMRD 839
 DB 781 SAYSGVLVTVYDLVTVLNLALISLGYFGPKCYMILFPERNTPAYFNSMIGYTMRD 839

RESULT 2
 ID 0920R7 PRELIMINARY; PRT; 843 AA.
 AC 0920R7;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Putative taste receptor tr2 (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H1STAR;
 RX MEDLINE=99159821; Pubmed=10052456;
 RA Hoon M.A., Adler E., Lindemeier J., Batley J.F., Ryba N.J.,
 RA Zuker C.S.;
 RT "Putative mammalian taste receptors: a class of taste-specific GPCRs
 RT with distinct topographic selectivity.";
 RL Cell 96:541-551(1999).
 DR EMBL: AF127390; AAD18070.1; -
 DR InterPro: IPR001828; ANF_receptor.

DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PRO0248; GPCRMRGR.
 DR PROSITE: PS00259; G_PROTEIN_RECPE_F3_4; 2.
 KW Receptor.
 FT NON_TER
 SQ SEQUENCE 843 AA; 95799 MW; D23AC22D21E049B8 CRC64;
 Query Match 72.1%; Score 3203.5; DB 11; Length 843;
 Best Local Similarity 70.6%; Pred. No. 1.3e-257;
 Matches 595; Conservative 108; Mismatches 133; Indels 7; Gaps 4;
 QY 1 MGPRAKTICSLFLFLVLAEP--AENSDFYLPDYLGLGFLSHAMKGIYHLNFGQVP 57
 DB 1 MGPRAKTICSLFLFLVLAEP--AENSDFYLPDYLGLGFLSHAMKGIYHLNFGQVP 57
 QY 58 MCKEYEVKVGYNIMQAMRAVEEINNDSSILPVLGLYELVYCYISNNVQPLVYFLAH 117
 DB 61 KNEFTKRVGLGYNIMQAMRAVEEINNCSSILPVLGLYELVYCYISNNVQPLVYFLAH 120
 QY 118 EDNLPLIOEDYSNYISRYAVVAVIGPDNSESVMYANFLSLFLPQITYSAISDELDRKVR 177
 DB 121 DDLLPLIKDYSQYMPRYAVVAVIGPDNSESATYSNIIHSLIPQITYSAISDELDRKVR 180
 QY 178 PALLRTTPSADHVEAMVQMLHFRNMWIIYVSSDTYGRDNGOLGERRAR--RDICIAF 236
 DB 181 PSMRLRTVPSATHHIEAMVQMLHFRNMWIIYVSSDTYGRDNGOLGERRAR--RDICIAF 240
 QY 237 QETPLQPNQNMNTSEERQRLVTVYDKLOOSTARVAVVESPDLTYHFENYLRONFGA 296
 DB 241 QETPLQPNQNMNTSEERQRLVTVYDKLOOSTARVAVVESPDLTYHFENYLRONFGA 300
 QY 297 VWIASEMAIDPVLANTELHGLTGLTIGTIOVPIPGFSEFEREMGPOAGPPPLSRTSOS 356
 DB 301 VWIASEMAIDPVLANTELHGLTGLTIGTIOVPIPGFSEFEREMGPOAGPPPLSRTSOS 360
 QY 357 YTCNQECNCLNATLSEFTIIRLSGERVYVYSVAVVAHALHSLGCDKSTCKRKYV 416
 DB 361 YTCNQECNCLNATLSEFTIIRLSGERVYVYSVAVVAHALHSLGCDKSTCKRKYV 420
 QY 417 PMOLLEIMKVNFTLDDHQIFPDQGVALHLEIYOMQDRSONFQSVASYPLQROLKNTOD 476
 DB 421 PMOLLEIMKVNFTLDDHQIFPDQGVALHLEIYOMQDRSONFQSVASYPLQROLKNTOD 480
 QY 477 NIODISMHTVNTIPMSCKRSKRGSGQKKKPGVJHVCCECICDLPCTFLNTEDEYECQ 536
 DB 481 YINNVSMTYTPRNTYVPMSCSKRGSGQKKKPGVJHVCCECICDLPCTFLNTEDEYECQ 540
 QY 537 ACPNNEMSYQSETCFQRQVLEHHEAPTAVALALGFLSTALIVYWRHQTIVSAG 596
 DB 541 ACPNNEMSYQSETCFQRQVLEHHEAPTAVALALGFLSTALIVYWRHQTIVSAG 600
 QY 597 RSAGGPMCFMLTLLVAVVAVVYVGPVKYSTCICRQALPPLCTICISCIARVFOIVCAF 656
 DB 601 RSAGGPMCFMLTLLVAVVAVVYVGPVKYSTCICRQALPPLCTICISCIARVFOIVCAF 660
 QY 657 CAFEMASRFPRAYSYVWRYOGPYVSMATYVVKWIVYIGMLARQOS--HPTDPDPKIX 714
 DB 661 CAFEMASRFPRAYSYVWRYOGPYVSMATYVVKWIVYIGMLARQOS--HPTDPDPKIX 719
 QY 715 TVISCNPNRNSLLEFNTSLDLLSVGFSFAVWKGELPTNYNEAKTITLSMPEYFSSV 774
 DB 720 TVISCNPNRNSLLEFNTSLDLLSVGFSFAVWKGELPTNYNEAKTITLSMPEYFSSV 779
 QY 775 LCTFMSASGVLYTVYDLVTVLNLALISLGYFGPKCYMILFPERNTPAYFNSMIGY 834
 DB 780 LCTFMSASGVLYTVYDLVTVLNLALISLGYFGPKCYMILFPERNTPAYFNSMIGY 839
 QY 835 MRR 837
 DB 840 MKR 842

RESULT 3

092514 PRELIMINARY: PRT: 843 AA.

AC 092514.

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Candidate taste receptor TIR2.

GN TASIR2 OR TIR2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J;

RX MEDLINE=21219400; PubMed=11319557;

RA Montmayeur J.P., Liberles S.D., Matsunami H., Buck L.B.;

RT "A candidate taste receptor gene near a sweet taste locus."

RL Nat. Neurosci. 4:492-498(2001).

DR EMBL: AF337041; AAK39438.1; -.

DR MGI: 1933546; Tasir2.

DR InterPro: IPR001828; ANF_receptor.

DR InterPro: IPR000337; GPCR_Mgr.

DR Pfam: PF00003; 7tm_3; 1.

DR Pfam: PF01094; ANF_receptor; 1.

DR PROSITE: PS50259; G_PROTEIN_RECIP_F3_4; 1.

KW Receptor.

SQ SEQUENCE 843 AA; 95735 MW; 0543834EA4E7AC2E CRC64;

Query Match 71.2%; Score 3163.5; DB 11; Length 843;

Best Local Similarity 69.2%; Pred. No. 2,8e-254;

Matches 583; Conservative 112; Mismatches 141; Indels 7; Gaps 4;

OY 1 MGPRKATICSFLFLMVAEP---AENSDFYLPDYLGLGFLSHANKGIYHLNFLOVP 57

DB 1 MGPRARTLHLFLHLALPKPVMALVGNDSDFHLAGDYLLGGLFTLHANKSVSHLSYLOVP 60

OY 58 MCKEYEVVIGNLMQANRFVVEEINNDSSLLPGVLGYEIVDYCYISNNQPVLYFLAH 117

DB 61 KCNEYNMVVLGNLMQANRFVVEEINNCSSLLPGVLGYEIVDYCYISNNQPGLYFLASQ 120

OY 118 EDNLPLIOEDYSNYISRYVAVIGPNSSESVMTANFLSLFLPOLYSAISDELBDKYRF 177

DB 121 IDDFLPILKDYSOYRPQYVAVIGPNSSESATVSNILSYFLVQYTYSAITDKLDDKRF 180

OY 178 PALRTPTSDHHEAVAMQOLMHRFNMNIIVLVSSDTYGRDNGOLLGERVARR-DICIAF 236

DB 181 PALMRTVPSAHTHIEAMQOLMHRFNMNIIVLVSSDDYGRNSHLLSQRLLTWTGICIAF 240

OY 237 OETLPTLPQNMNTSEERORLVTYDKLQOSTARVYVPSPLTIYHFENEVLRNFTGCA 296

DB 241 OEVLPVPEPNOAVRPEEDODLNDLKLRTSARVAVIFSPSLSHNFFREVLARNFTGF 300

OY 297 VWASESAIDPVNLHNLTELGLTFTLGIITQSVDPISGFSEERWGPQAGPPLSRTSOS 356

DB 301 VWASESAIDPVNLHNLTELRLRTGTFTLGIQSVSIPGSGQRYVRHNDPEYMPMETSLR 360

OY 357 YTCNOCNCLNATLSFTTILRLSGERYVYSYSAVYAAHLSHLGCDKSTCKRKYV 416

DB 361 TTCNODCCACAMNTESFNNVLMLSGERVYYSYSAVYAAHTLHRLHLCNOCVRCYKOIY 420

OY 417 PMQLLEIMKVFETLDDIIFDPDQGVALHLEIYQMDRSONFOSVASYSPLOROLK 476

DB 421 PMQLLEIMKVFETLDDIIFDPDQGVALHLEIYQMDRSONFOSVASYSPLETRLT 480

OY 477 NIQDISMHTVNTIPMSCKRCSGQKKRPVIGVCCFECTIDCLPGLTNTEDEYEQ 536

DB 481 YISNYSWTPNNTVPIISCKSCQCGQMKRPGLHPCFECVDCPGYILNNSVDEPNCL 540

OY 537 ACPNNEKSYQSTSCFKQOLVLEHHEAPTIVALLALGISTALIVYTRHHTPIV 596

DB 541 ACPNNEKSYQSTSCFKQOLVLEHHEAPTIVALLALGISTALIVYTRHHTPIV 596

RESULT 4

092338 PRELIMINARY: PRT: 843 AA.

AC 092338.

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Sweet taste receptor TIR2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC Nelson G., Hoon M.A., Chandrasekar J., Zhang Y., Ryba N.J.P., Zuker C.S.;

RT "Mammalian Sweet Taste Receptors."

RL Cell 0:0-0(2001).

DR EMBL: AY032623; AAK51604.1; -.

DR InterPro: IPR001828; ANF_receptor.

DR InterPro: IPR000337; GPCR_Mgr.

DR Pfam: PF00003; 7tm_3; 1.

DR Pfam: PF01094; ANF_receptor; 1.

DR PROSITE: PS50259; G_PROTEIN_RECIP_F3_4; 1.

KW Receptor.

SQ SEQUENCE 843 AA; 95752 MW; FD3C6B9B23270DA3 CRC64;

Query Match 70.9%; Score 3151.5; DB 11; Length 843;

Best Local Similarity 68.9%; Pred. No. 2,8e-253;

Matches 581; Conservative 113; Mismatches 142; Indels 7; Gaps 4;

OY 1 MGPRKATICSFLFLMVAEP---AENSDFYLPDYLGLGFLSHANKGIYHLNFLOVP 57

DB 1 MGPRARTLHLFLHLALPKPVMALVGNDSDFHLAGDYLLGGLFTLHANKSVSHLSYLOVP 60

OY 58 MCKEYEVVIGNLMQANRFVVEEINNDSSLLPGVLGYEIVDYCYISNNQPVLYFLAH 117

DB 61 KCNEYNMVVLGNLMQANRFVVEEINNCSSLLPGVLGYEIVDYCYISNNQPGLYFLASQ 120

OY 118 EDNLPLIOEDYSNYISRYVAVIGPNSSESVMTANFLSLFLPOLYSAISDELBDKYRF 177

DB 121 IDDFLPILKDYSOYRPQYVAVIGPNSSESATVSNILSYFLVQYTYSAITDKLDDKRF 180

OY 178 PALRTPTSDHHEAVAMQOLMHRFNMNIIVLVSSDTYGRDNGOLLGERVARR-DICIAF 236

DB 181 PALMRTVPSAHTHIEAMQOLMHRFNMNIIVLVSSDDYGRNSHLLSQRLLTWTGICIAF 240

OY 237 OETLPTLPQNMNTSEERORLVTYDKLQOSTARVYVPSPLTIYHFENEVLRNFTGCA 296

DB 241 OEVLPVPEPNOAVRPEEDODLNDLKLRTSARVAVIFSPSLSHNFFREVLARNFTGF 300

OY 541 SCPSGMSYKNNIACFKRRLAFLHHEWPTIVYTIILALGISTALILIFMRHQPMPV 600

OY 597 RSAGGPMCFIMTLVVAVMVVYVGPVKSTCLCRALPFCPTICISIAVBSFOV 656

DB 601 RSAGGPMCFIMTLVVAVMVVYVGPVKSTCLCRALPFCPTICISIAVBSFOV 660

OY 657 CAFKASRFRPRASYWVRQGPVSMATFTVKMIVVIGMLARPQSHP--RTDDDPKI 714

DB 661 CYFKARRLPSAYGFMHRHGPYFVAITAKVALVAGNMMLA-TTINIGITDDDPRI 719

OY 715 TIVSCPNYRNSLFTNSLIDLLSVGFSFAYMGKELPTNNEAKFTILSMFTYTSVS 774

DB 720 ILSCHPNYRNSLFTNSLIDLLSVGFSFAYMGKELPTNNEAKFTILSMFTYTSVS 779

OY 775 ICTFMSASGVLTIVDLYVNLALISLGFGRKCHILFEPERNPAYNSMIQGT 834

DB 780 ICTFMSVHDGVLTIVDLYVNLALISLGFGRKCHILFEPERNPAYNSMIQGT 839

OY 835 MRR 837

DB 840 MRR 842

QY	297	VWASEMAIDPVLAHNLTELGHSGLGITIQSPVPEPSEFRMGQAGPPLSPRISQS	355
QY	297	VWASEMAIDPVLAHNLTELGHSGLGITIQSPVPEPSEFRMGQAGPPLSPRISQS	355
Db	301	VWASEMAIDPVLAHNLTELGHSGLGITIQSPVPEPSEFRMGQAGPPLSPRISQS	366
QY	357	YTCNDECONCINATLSFTFIIIRLSEGERVVSVAAYVAVALHSLLGCDSTCKRVVY	418
Db	361	YTCNDECONCINATLSFTFIIIRLSEGERVVSVAAYVAVALHSLLGCDSTCKRVVY	422
QY	417	PMQLLEETAKVYVFTLLDQIFEPDQGVALLLEYVOMQDRSNPFQSVASTYLQRLK	478
QY	421	PMQLLEETAKVYVFTLLDQIFEPDQGVALLLEYVOMQDRSNPFQSVASTYLQRLK	480
QY	477	NIDISMTVNTVNTIPMSCKSCGOKKRPVGIHVCCFECIDLPFTLNHTDEYEQ	536
Db	481	YISNVSWTTPNNTVYISMSCSQCPGOKKRTIGLHPCCFECYDPPDTYLRNSYDERNCL	548
QY	537	ACPNNMWSYQSESTCFKRLQVLEWHEAPRTAVALALGELSTLALVIFWRHPTPIV	596
Db	541	SCPQSMWSYKNNINACFKRLALFLEWHEPTVITVITLALGELSTLALVIFWRHPTPIV	600
QY	597	RSAGPMPCLMTLLVLAIVMVVPTVGPVKYSTCLQRAFLPCLFTICISCIASVRSQIV	656
Db	601	RSAGPMPCLMTLLVLAIVMVVPTVGPVKYSTCLQRAFLPCLFTICISCIASVRSQIV	660
QY	657	CAFKASAFPRPAYSVWVYOCQPVYSMAITLKNVYVIGLARQSH--RTCPDDPKI	714
Db	661	CAFKASAFPRPAYSVWVYOCQPVYSMAITLKNVYVIGLARQSH--RTCPDDPKI	718
QY	715	TVSCNPYRNSLLENTSIDLLSVYGFSPAYMGKELPTNNEAKFTLSMTFFYTSVS	774
Db	720	ILTSCHPNYRNGLENTSMOILLSTVIGRSPAYMGKELPTNNEAKFTLSMTFFYTSVS	778
QY	775	ICTFMSAIVSYLVTVOLLVYVNLALSLGTFPGPKMILFYDEBRTPAFNSMIOGYT	833
Db	780	ICTFMSAIVSYLVTVOLLVYVNLALSLGTFPGPKMILFYDEBRTPAFNSMIOGYT	838
QY	835	MRR 837	
Db	840	MRR 842	
RESULT 5			
Q920R8			
ID	Q920R8	PRELIMINARY;	PRT; 840 AA.
AC	Q920R8		
DT	01-MAY-1999	(TREMBLrel. 10, Created)	
DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)	
DT	01-JUN-2002	(TREMBLrel. 21, last annotation update)	
DE	Putative taste receptor TRL (Fragment).		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus		
OX	NCBI_TaxID=10116;		
RM	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-MISTAR;		
RX	MEDLINE=99159821; PubMed=10052456;		
RA	Hoon M.A., Adler E., Lindemeyer J., Battey J.F., Ryba N.J.,		
RA	Zuker C.S.;		
RT	"Putative mammalian taste receptors: a class of taste-specific GPCRs		
RT	with distinct topographic selectivity.";		
RL	Cell 96:541-551(1999).		
DR	EMBL; AF127389; A018069.1; -		
DR	InterPro; IPR001828; ANF_Mgcr.		
DR	InterPro; IPR000337; GPCR_Tgcr.		
DR	Pfam; PF00003; 7tm_3; 1.		
DR	Pfam; PF01094; ANF_receptor; 1.		
DR	PRINTS; PR00248; GPCRMR.		
DR	PROSITE; PS50259; G_PROTEIN_RCRP_F3_4; 2.		
FW	Receptor.		
FT	NON_TER	840	840
SO	SEQUENCE	840 AA;	93496 MW; 1FCFB7EFC6BB45DB CRC64;

Query Match	32.7%	Score 1454	DB 11	Length 840
Best Local Similarity	39.5%	Pred. No. 4.5e-112		
Matches 325	Conservative 123	Mismatches 346	Indels 28	Gaps
QY	27	FYLPDDYLGGEGFSTSHAMMKGIIVHNFQVPMC--KEVEVKYIGYMLQAMRFAVEEINN	85	
Db	31	FELPDQFLLAGFLSHGOCIGYRHNPL--VTSQDRPDSFNGHYTLQPMAFYEEINS	88	
QY	86	SSLDFVLLGIEIVDYCTISNNVQVLYFLAHE-DNLLPIQEDYSNTISRVAAYIGDNS	144	
Db	89	SALLFENITLGYELXYDVCSESANVYATLLRYLALOGRRHIEIQKDINHSSKVAFIGPNT	146	
QY	145	EAMVIVANFLSFLLPQITYSISGELBDKRFVLTTPPSADHNVAMQOLMHFPMN	204	
Db	149	DNAVTTALLGLFPLMLPYLSYEAASSVYLSAKRRFSEFLTPSDRHOVYVMQOLSGW	208	
QY	205	WIVIVSSPTYRDNQGLGIEKRVARDICIAFOEFLPILOPNQMTSEEROLVYIDKL	266	
Db	209	WISLISYGDYQDLOVQALIELAYRGICVAFKDYLP-----SARGDGRMSQMHL	266	
QY	265	QOSTARVVVSPDILTYHFENEVLKQNFQAVWTASSMAIDPYLNLTELGHGLFVG	322	
Db	263	AGARTTYYVVFENRHIAREFRSVYLAITGKVVWASDEMAISYITGVGIGIGYVG	322	
QY	325	ITIGSPVPIGSEFEREMGRQAPPLSTQSUTC--NOEDNCNLAPLSTNTILRLSGE	388	
Db	323	VAVQORQVPGLEKEFEESYVRAVTAAPSACPESSWCSTQOLRECTFTTRNPTGAFSM	388	
QY	383	RVVYSVSAVYVAVAHLSLGLCDSTCKRVVYPMQLLEIWKNTFLDHOIFPDQ	444	
Db	383	SAAYAEVAVYVAVAGHQLLGGTSEICSRGVYPMQLQYKYNFLHNTVAFDNG	444	
QY	443	DVALHLEIYQOMDRSQNPQSV--ASTYPLQRLKNIQDISMTVNTITMSCKSRQ	500	
Db	443	DVLGYDDIAMDQNGPEMTFELIGSASLSPVHDI-NTKTIQMGKNQVPSVCTDCL	500	
QY	501	SGOKKRPVGIHVCEGIDCLPGFLNTEDEYECQAPNNMSYQSEFCEKROLVLE	566	
Db	502	AGHHRVYVGSHHCCCEYCPCENGTFLNMS-LLHCQPCGTBEMAKESTCPRVYELA	566	
QY	561	WHEAPITIAVALLAALGFSTIAIIV-----IFMRHPTPIVRSAGGPMCEMLTLLVA	614	
Db	561	WHE--PISLVLIA-----NTLLLLLVTAGLFAHNFHTPVVRSAGGRCLMLGSLVAG	614	
QY	615	YAVVYVYVYVPPVYVSCLRQALFPLCTIICISIAVRSFQYCAFKMASRRPRAYSTVVR	677	
Db	615	SCSFYSFGEPTVPACLLKQPLFSIGFAIFLSCILIRSFOLVITFEFSTKVPYRTYTAQ	677	
QY	675	YQGPVYSMAFIVLWMAIVYIGMLARPQSHPTDDDKRITIVSCNPVYRNSLLEFNTSLD	734	
Db	675	NHGAGLFVYSVTVALLICLTWLVMTTRPPRTREYRPHALVILECTEVNSGFLAFTHN	734	
QY	735	LLLSVGFSAFMAKGLDPLTNNEAFTILSMTFYTTSSVSLCTFMSAYSGVLYIVDLLV	794	
Db	735	LLLSITFVCSLGLGELDENYEAACVFFSLLLNFSVMIAFETMASIYQGYLPAVNVLA	794	
QY	795	TYLNLALSLGFGKCYMILFYPRRNPATFNSMIGQYTKR	836	
Db	795	GLTTLGGFSGYFLPKCYVILCPRLPNTNEHQASIDQYTRR	836	
RESULT 6				
ID	092515	PRELIMINARY	PRF	842 AA.
AC	092515			
DT	01-DEC-2001	(TREMBLrel. 19, created)		
DT	01-DEC-2001	(TREMBLrel. 19, last sequence update)		
DE	01-JUN-2002	(TREMBLrel. 21, last annotation update)		
DR		candidate taste receptor TIR1.		
GN		TASIR1 OR TIR1.		
OC		Mus musculus (Mouse).		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=C219400; PubMed=11319557;
 RA Montmayeur J.P., Liberles S.D., Matsunami H., Buck L.B.,
 "A candidate taste receptor gene near a sweet taste locus,"
 RL Nat. Neurosci. 4:492-498(2001).
 DR EMBL; AF337040; AK39437.1; .
 DR MGI; MGI:1927505; Tas1r1.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000345; CyTC_heme_bind.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS50259; G_PROTEIN_RECER_F3_4; 1.
 DR Receptor.
 KW Receptor.
 SQ SEQUENCE 842 AA; 93471 MW; FDBA0CC1BD45DEB7 CRC64;

Query Match 32.4%; Score 1441; DB 11; Length 842;
 Best Local Similarity 39.7%; Pred. No. 5.4e-111;
 Matches 328; Conservative 123; Mismatches 339; Indels 36; Gaps 15;

OY 27 FYLPDGYLLGGLFSLHANKKGIHNLFLQVPMC-KEYEYKVIQYINLQMAFRVAEINND 85
 DB 33 FSLPDPFLLAGLFSLHADLOVRHRL--VTSCHSDSDSNGHGHLFQMRRTVEINNS 90
 OY 86 SLLPGVLLGEIVDVCYISNNVOPVLYFLAHE-DNLPLIOEDYSNYISRYVAVIGPDS 144
 DB 91 TALLPNITIGLYELDYVCSSESNVYATLRVLAQOGHLEMDLRNHSKVVALLGPDT 150
 OY 145 ESMVTVANFLSLFLPQITTSASIDELDKVRFALLRTTPSADHVEAMVOLMLHFRN 204
 DB 151 DHAVTTAALLSPFLMPLVSEASVYLSGKRKFPFLRTIPSDKQVEIVTLQSGFV 210
 OY 205 WTIIVSSDTVGRDNGQLGGEVARRDICIAFOETPLTQPNQNTSEEROLVITVDK 264
 DB 211 WISLVSGIDGQGLVQALELATPRGICVAFKDVPL---SAQAGDPRMOMML---RL 264
 OY 265 QOSTARVVVSPDLTLVFFENEVLQRNFTGAVWIASSEMAIDPVLNLTGLHGLTFGL 324
 DB 265 AARFTVVVVFENRHLAGFFRSVLANLTGKWIASEDMAISTYITNVPGIGIGTVLG 324
 OY 325 ITIQSVPIPGSEFEREMRQA--GPPPLSRTSQSTYCNQECNCLNATISFTIIRLSE 382
 DB 325 VAIQQRQVPGGLKEFESESYQAVMGAPRTCPGSGMCTNOLCRECHAFITWNNPBLGAFSM 384
 OY 383 RYVSVSVSAVYVAHALSLIGCDKSTCTKRVVYPMOLLEITWKYNFTLLDQIFPDPOG 442
 DB 385 SAAYVAVYEVAVVAGHQLGCTGCTCARGVYPMOLLOQYIKYNFLHKTVAFPDQK 444
 OY 443 DVALLLEIVQWQDRSQNPQSV--ASYPIQORQLKNIQDISMHTVNTTPSKMSKRCQ 500
 DB 445 DPLGYDDIIMANGPEWTFEYIGSASLSPVHLDI-NKTKIQMHGKNQVAVSVCTRDL 503
 OY 501 SGQKKRPVGHVCECTIDCLPGFTLNTHDEYEGOACPNMWSQSESCRKQDLVLE 560
 DB 504 ESHHNLVMSHRCFECMECEAGTFLN-TSELHTQPCGTEWMAPEGSSACSRVEEFG 562
 OY 561 WHEAPTLIAVALAAGLSTALIVF-----NRHQTPIVRSAGGPMCLMTILLY 613
 DB 563 WHE--PISIVLLAA-----NTLLILLIGAGLFAFR-LHTPVYRBSAGGLCTLMGLSLVA 615
 OY 614 AYVVVYVYVGPVSTCLQALPPLCTTICISIAVRSFOIVCAFKNASRPRAVSYV 673
 DB 616 GSCSYSEFGKPTVPACILRPLFSLGFAIFLSCLTIRSFOLVITFKSTKVPYTHWA 675
 OY 674 RQSGPYVSAFTVLYKMYIVVIGMLARQSHRTDPD---DKITIVSCNPMYKRSLEFN 730
 DB 676 QNHGAGI---FVIVSVTHVFLCTLWLANMTRPTREYQRFHVLILECTEVNSVGFVA 732

OY 731 TSLDILLVYSGFAYMGKELPTNYNEAKFITLSMTEFTSSVSLQTEMSAYGLVNTY 790
 DB 733 FAHNILLISTVCSYGLGELPENTNEAKCYFSLHFPVSIATFTMSITQSTLPAV 792
 OY 791 DLLVTVNLNLAISLGYFGRCYMITLFYPERNPVAFNSMIGCYTKR 836
 DB 793 NVLAGLATLNSGSGFLPKCYVILLCRPELNTEHFRQASIQDYTKR 838

RESULT 7

O99PG5 PRELIMINARY; PRT; 842 AA.
 AC O99PG5;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Putative sweet taste receptor TIR1 (Fragment).
 GN TASIR1 OR GPR70.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129P3/J;
 RX MEDLINE=21030739; PubMed=11178737;
 RA Li X., Inoue M., Reed D.R., Hugue T., Puchalski R.B., Tordoff M.G.,
 Nihomiya Y., Beauchamp G.K., Bachmanov A.A.;
 "High-resolution genetic mapping of the saccharin preference locus
 (Sac) and the putative sweet taste receptor (TIR1) gene (Gpr70) to
 RT mouse distal chromosome 4";
 RL Mamm. Genome 12:13-16(2001).
 DR EMBL; AF301162; AK07092.1; .
 DR MGI; MGI:1927505; Tas1r1.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000345; CyTC_heme_bind.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS50259; G_PROTEIN_RECER_F3_4; 1.
 DR Receptor.
 KW NON_TER.
 SQ SEQUENCE 842 AA; 93428 MW; 39739A2FF482D33F CRC64;

Query Match 32.4%; Score 1440; DB 11; Length 842;
 Best Local Similarity 39.7%; Pred. No. 6.6e-111;
 Matches 328; Conservative 122; Mismatches 340; Indels 36; Gaps 15;

OY 27 FYLPDGYLLGGLFSLHANKKGIHNLFLQVPMC-KEYEYKVIQYINLQMAFRVAEINND 85
 DB 33 FSLPDPFLLAGLFSLHADLOVRHRL--VTSCHSDSDSNGHGHLFQMRRTVEINNS 90
 OY 86 SLLPGVLLGEIVDVCYISNNVOPVLYFLAHE-DNLPLIOEDYSNYISRYVAVIGPDS 144
 DB 91 TALLPNITIGLYELDYVCSSESNVYATLRVLAQOGHLEMDLRNHSKVVALLGPDT 150
 OY 145 ESMVTVANFLSLFLPQITTSASIDELDKVRFALLRTTPSADHVEAMVOLMLHFRN 204
 DB 151 DHAVTTAALLSPFLMPLVSEASVYLSGKRKFPFLRTIPSDKQVEIVTLQSGFV 210
 OY 205 WTIIVSSDTVGRDNGQLGGEVARRDICIAFOETPLTQPNQNTSEEROLVITVDK 264
 DB 211 WISLVSGIDGQGLVQALELATPRGICVAFKDVPL---SAQAGDPRMOMML---RL 264
 OY 265 QOSTARVVVSPDLTLVFFENEVLQRNFTGAVWIASSEMAIDPVLNLTGLHGLTFGL 324
 DB 265 AARFTVVVVFENRHLAGFFRSVLANLTGKWIASEDMAISTYITNVPGIGIGTVLG 324
 OY 325 ITIQSVPIPGSEFEREMRQA--GPPPLSRTSQSTYCNQECNCLNATISFTIIRLSE 382
 DB 325 VAIQQRQVPGGLKEFESESYQAVMGAPRTCPGSGMCTNOLCRECHAFITWNNPBLGAFSM 384

QY 383 RYVSVYSAVYAAVAHSHLGGCDKSTCRKRVYPMOLLEIMKVNFTLLDHOIFPDG 442
 DB 385 SAAYNYEAYVAAGHGLHOLGCTSGCARGPYPMOLLOQITKVNFLHKKTYAFDNG 444
 QY 443 DVALHLEIVOMDRSQNPQSV--ASYPLQRLKNIODISMTVNTTIPMSCKSRQ 500
 DB 445 DPLGYDYDIAMDWNGPMTREVEIGSASLSPVHDI-NKRTIOWHGKNNQVPVSVCTRDCL 503
 QY 501 SGQKKRKGIVHCCFECIDLPGLFNLHTEDEYECACPNMNSYQSECTCFKRLVLE 560
 DB 504 EGHRLVMGSHCCFECMPPEAGTFLN-TSELHQCQCGTEENAPBESSACFSRTVEFLG 562
 QY 561 WHEAPITVAVALAIGLSTLAIIVF-----WRHFQPIVRSAGCPMCFMLTLLV 613
 DB 563 WHE--PISVLIAA-----NTLLLLLLIGTAGLFAMR-LHTPVRSAGRCFLMLGSLVA 615
 QY 614 AYAVPYVYPPRYVSTCLCQALFPLCTTCISCIASRSQIYCAEFMASRPPRAYSVW 673
 DB 616 GSCSLXSFEGEPVPACILRQPLFSLGFAFLSCLTIRSFQLYIFEFSTRKVPFFYHTWA 675
 QY 674 RYGPVYMAFITVLMKVIIVIGMLARQSHPRDPP--DKPTIYSCPNRNSLTFN 730
 DB 676 QNHGAGI--FVIYSSIVHFLCLTTLAMTPRPTREYQRPPLVILCETEVNSVGLVA 732
 QY 731 TSDLLLSYVGFSAVYKGLPTNYNEAKFTLSMPTFTSSVSLCTFMSAGVLTIV 790
 DB 733 FAHNLILSTIFVCSYLGKELPENYNAKCVTSILHFVSMIAFFMSSIIYQSYLPAY 792
 QY 791 DLVTVYVNLALAIISGTRGPKCYMILFPERNTPAYFNSMIOGTMR 836
 DB 793 NVLAGLNTLSGFGSYFLPKCYVILCRPELNTNTEHFQASIODYTRR 838

RESULT 8

0923J9

ID 0923J9 PRELIMINARY; PRT; 842 AA.

AC 0923J9;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE Taste receptor TIRI.
 GN TASIR1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nelson G., Hoon M.A., Chandrasekar J., Zhang Y., Ryba N.J.P.,
 RA Zuker C.S.,
 RT "Mammalian Sweet Taste Receptors";
 RL Cell 0:0-0(2001).
 DR EMBL: AY032622; AAK51603.1; -
 DR MGD: MGI:1927505; Tasir1.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm.3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS50259; G_PROTEIN_RECEP_F3_4; 1.
 KW Receptor.
 SQ SEQUENCE 842 AA; 93455 MW; D4DCEB0959E991A CRC64;

Query Match

Best Local Similarity 32.3%; Score 1434; DB 11; Length 842;
 Matches 327; Conservativity 123; Mismatches 340; Indels 36; Gaps 15;

QY 27 FYVPGVILAGLGSLSHNMKGIVLNFLOVPMG-KEYEVVIGVNLMOARPAVEELND 85
 DB 33 FSLPGDFLAGLSLHADCLOVRRRL--VTSCDRSDSFGHGHIHLQAMRFVEELINS 90

QY 86 SLLPGVILAGELYDYVYCISSNNYQVPLYFLAHE-DNLLPIQEDYSNYSIRVAVIGDNS 144
 DB 91 TALLPNTLTGELYDYVCSSESNAYATLRVAQOCTGLEMQRLRNHSSKVALIGPNT 150
 QY 145 ESWTVANFSLLELLPITTSALSDELKDYRRPALLKRTPPSDHRYEAVQMLHRMN 204
 DB 151 DHAVTTALLSPFLMPLVSYEASVILSGRKRPSPFLRTIPSKYQYQEVIRLLQSGW 210
 QY 205 WITVLSDDYGRNGOLLEERVARDICIAFOETLTLPQNNMTSEEQRLVITYDKL 264
 DB 211 WISVSYGTYGQGLVQALBELATPRGICVAFKDYVL-----SAQAGPROMRM--RL 264
 QY 265 QOSTARVYVESPDLLYHFNELRQNTGAVWIASESAIDPVLNLTGLHGTFLG 324
 DB 265 ARATTVVYVVSNNHLAGVEFRSVLANLGKVIASEDAISYITNVPGIOGIVLG 324
 QY 325 ITTQSVPIPGFSEREPGPA--GPPLSRTSOSYTQNOECDCLNATLSFNILRSGE 382
 DB 325 VALQORQVPLKKEESYQAVMGAPRTCPGSGWCTNOLCRCHAFPTTNMBELGAFSM 384
 QY 383 RYVSVYSAVYAAVAHSHLGGCDKSTCRKRVYPMOLLEIMKVNFTLLDHOIFPDG 442
 DB 385 SAAYNYEAYVAAGHGLHOLGCTSGCARGPYPMOLLOQITKVNFLHKKTYAFDNG 444
 QY 443 DVALHLEIVOMDRSQNPQSV--ASYPLQRLKNIODISMTVNTTIPMSCKSRQ 500
 DB 445 DPLGYDYDIAMDWNGPMTREVEIGSASLSPVHDI-NKRTIOWHGKNNQVPVSVCTRDCL 503
 QY 501 SGQKKRKGIVHCCFECIDLPGLFNLHTEDEYECACPNMNSYQSECTCFKRLVLE 560
 DB 504 EGHRLVMGSHCCFECMPPEAGTFLN-TSELHQCQCGTEENAPBESSACFSRTVEFLG 562
 QY 561 WHEAPITVAVALAIGLSTLAIIVF-----WRHFQPIVRSAGCPMCFMLTLLV 613
 DB 563 WHE--PISVLIAA-----NTLLLLLLIGTAGLFAMR-LHTPVRSAGRCFLMLGSLVA 615
 QY 614 AYAVPYVYPPRYVSTCLCQALFPLCTTCISCIASRSQIYCAEFMASRPPRAYSVW 673
 DB 616 GSCSLXSFEGEPVPACILRQPLFSLGFAFLSCLTIRSFQLYIFEFSTRKVPFFYHTWA 675
 QY 674 RYGPVYMAFITVLMKVIIVIGMLARQSHPRDPP--DKPTIYSCPNRNSLTFN 730
 DB 676 QNHGAGI--FVIYSSIVHFLCLTTLAMTPRPTREYQRPPLVILCETEVNSVGLVA 732
 QY 731 TSDLLLSYVGFSAVYKGLPTNYNEAKFTLSMPTFTSSVSLCTFMSAGVLTIV 790
 DB 733 FAHNLILSTIFVCSYLGKELPENYNAKCVTSILHFVSMIAFFMSSIIYQSYLPAY 792
 QY 791 DLVTVYVNLALAIISGTRGPKCYMILFPERNTPAYFNSMIOGTMR 836
 DB 793 NVLAGLNTLSGFGSYFLPKCYVILCRPELNTNTEHFQASIODYTRR 838

RESULT 9

099PG6

ID 099PG6 PRELIMINARY; PRT; 842 AA.

AC 099PG6;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, last annotation update)
 DE Putative sweet taste receptor TIRI (Fragment).
 GN TASIR1 OR GPR70.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=C57BL/6BYJ;
 RC MEDLINE=21030739; PubMed=11178737;
 RA Li X., Inoue M., Reed D.R., Huque T., Puchalski R.B., Tordoff M.G.,
 RA Ntouniyla Y., Beauchamp G.K., Bachmanov A.A.;
 RT "High-resolution genetic mapping of the saccharin preference locus

RT (Sac) and the putative sweet taste receptor (TLRI) gene (Gpr70) to
 RT mouse distal chromosome 4.1;
 RL Mamm. Genome 12:13-16(2001).
 DR EMBL: AF301161; AK07091.1;
 DR MGI:1927505; Tas1r1.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS50259; G_PROTEIN_RECPT_F3_4; 1.
 KW Receptor.
 FT NON_TER
 SO SEQUENCE 842 AA; 93425 MW; 58826C43F5DD352B CRC64;
 Query Match 32.2%; Score 1431; DB 11; Length 842;
 Best Local Similarity 39.6%; Pred. No. 3.7e-110;
 Matches 327; Conservative 124; Mismatches 339; Indels 36; Gaps 15;
 QY 27 FYLDGDIYLGFLSHAMKGIHNLQVPMG-KEYEVKVIQVLMQAMRAVEIIND 85
 DB 33 FSLPGDFLLAGLFSIHADCLQVRHRL--VTSRSDSFNGHYHLQAMFEVVEINNS 90
 QY 86 SSLPGVLLGEIVDVCYISNNQPVLYFLAHE-DNLLPIQEDYSNYSRYVAVVGPDNS 144
 DB 91 TALPNTLIGELIDVCESSNVATLRYLAQOGTGLEMORDLNHSKVALGLPNT 150
 QY 145 ESMVTANFLSLFLPQTYAISDELNDKRYFPALLRTSPADHIVAMVQLMHPMN 204
 DB 151 DHAVTATALLSPFLMPLVSYEASSVILSGKRKFPFLRTISDKQVEIVIRLLQSPGMV 210
 QY 205 WIIYVSSDTYGRDNGQLLGERVARRDICIAFOETLPTLPQNPQNTSEEROLVTYDKL 264
 DB 211 WISLVGSYGDQGLQVQALAEELAPRGICVAFKNVPL--SAQAGDPRMOMML--RL 264
 QY 265 QOSTARVVVSPDLTLYHFENEVLRONFTGAVMIASEMAIDPVLNLTGHLGFTGL 324
 DB 265 ARARTVTVVNSNRHLDGVFRSVYLANLTGKVMIASEDMAISTIRPVSIGIGITVVG 324
 QY 325 ITIOSVPIGSEFEREMPOA--GPPPLSRTSOSTYCNQEDNCINATLSFNTILRSGE 382
 DB 325 VAIOORQVPGKEEESYVQAVMGAPRTCPESGMCNQLRCHAFPTWMPDELGAISM 384
 QY 383 RVVSVYSAVVAVAHAHSLGCDKSTCKRVYIPWQLEELWKNFLLDHOIFPDQG 442
 DB 385 SAAYNVYAVVAVAHGLHQLGCTSGTCARGVHPWQLQOIKYKNFLHKKVAFDQK 444
 QY 443 DVALLLEIVQOMDRSONPQSV--ASYYPLOROLKNIDISMHTVNNTPMSMCKRQ 500
 DB 445 DPLGIYDIANDMNGPEPTFEVIGSASLSPVHLDI-NKTKIOMHCKNOQVSVCTRCL 503
 QY 501 SGOKRKPVGIHVCCFECIDCLPGFLNTEDEYECQACPNMWSYOSTSECFKROLVLE 560
 DB 504 EGHNLVYVSHHCCEPCPCAGFLN-TSELHTQPCGTEMAEAGSSACFSRVEELG 562
 QY 561 WHEAPTVALLAALGFLSTAILVIF-----WRHPOTPIVSSAGPQCFMLTLLLV 613
 DB 563 WHE--PISLVLLAA---NTLLLLLLITAGLFAWR-LHTFVVRSGARLFLMIGSLVYA 615
 QY 614 AYVVVYVVGPPKYSTCICRQALPFLCTICISCIASVSPQIVCAFKNASRPRAYSVW 673
 DB 616 GSCSIYSFGRVYPACILRPLSLGFAIPLSCLTISFDVLVILFKSTVPTFYHTWA 675
 QY 674 RYOGPVYMAFIVLKAVIVIGMLARPOSHRTPD--DPKITIVGCNPYRSLFN 730
 DB 676 QNHGAGI--FYIVSSYVHFLCLTWLMMTPRPTREYQRPHVLILCTEVSNGFLVA 732
 QY 731 TSLDILLVSGFAYMKEKLPNTNNEAKFTILSKTFYFTSSVSLCTMSAYSGVLVTV 790
 DB 733 FAHNILLISITFVCSYLKELPENYNEAKVTFSLHIFVSMIAFTMSSITQOGLYDPAV 792

QY 791 DLTVVLMALISLGYRPGKCMILFPERNTPAFNSMIGQYMR 836
 DB 793 NVLAGLATLGGSGYFLPKCYVILCRPELNTEHFQASIDQYTR 838
 RESULT 10
 Q8TDJ9
 ID Q8TDJ9 PRELIMINARY; PRT; 763 AA.
 AC Q8TDJ9;
 DT 01-JUN-2002 (TREMblrel, 21, Created)
 DT 01-JUN-2002 (TREMblrel, 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel, 21, Last annotation update)
 DE Gm148 form B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21888635; PubMed=11891061;
 RA Makalowska I., Scod R., Faruque M.U., Hu P., Robbins C.M.,
 RA Eddings E.M., Mestre J.D., Baxevanis A.D., Carpten J.D.;
 RT Identification of six novel genes by experimental validation of
 RT GeneMachine predicted genes.
 RL Gene 284:203-213(2002).
 DR EMBL: AF387618; AAL91359.1;
 SO SEQUENCE 763 AA; 84427 MW; 593616575D6BD17D CRC64;
 Query Match 28.2%; Score 1253.5; DB 4; Length 763;
 Best Local Similarity 36.9%; Pred. No. 1.9e-95;
 Matches 287; Conservative 124; Mismatches 332; Indels 35; Gaps 12;
 QY 75 MRAVEEINNDSSLPGVLLGEIVDVCYISNNQPVLYFLA---HEDNLLPIQEDYSN 130
 DB 1 MRLGVEEINNSTALLPNTLIGYOLVDCSDSANVYATRLVSLPQNH---IELQGLIH 57
 QY 131 YISRVAVYIPDNSESVTVANFLSLFLPQTYAISDELNDKRYFPALLRTSPADH 190
 DB 58 YSPYLVAYIGDSTRATTAALSLPFLVPMIYASSETLSYKQYFSPLRTINDKQ 117
 QY 191 VEAMVQMLHFRMNIIVLVSSDTYGRDNGQLLGERVARRDICIAFOETLPTLPQNP 249
 DB 118 VETWVLLQKFGWVWISLVGSSDDYQGLQVQALENQATGQGICAFKIDMPSAQVD- 175
 QY 250 TSEEROLVTYDKQOSTARVVYFSPDLTLYHFENEVLRONFTGAVMIASEMAIDPV 309
 DB 176 ----ERMQCLMRHLAQAGATVVVFSRQALARVEFESVLTNLGKVMVASEVALSRH 230
 QY 310 LHNLTGHLGFTGLGITIOSVPIPGSEFEREMPOA--GPPPLSRTSOSTYCNQEDNC 366
 DB 231 ITGVGIGIRIGVNLGVALQKRAVPGIKAFEEYARADKXARPCKSGMC--SSNOLCRQC 289
 QY 367 LNAITLSENTILRLSGERVYVSYSAVVAVAHAHSLGCDKSTCKRVYIPWQLEELW 426
 DB 290 QAFMAHTPKLKAFASSNAVYAVVAHGLHQLLCAAGACSGRGVYPMQLEQJHK 349
 QY 427 VNFITLDHOIFPDQDVALHLEIYQOMDRSONPQSVAS--YTPLOROLKNIDISM 484
 DB 350 VHFLLHKDVAFENDRDLSSYNIITAMWNGKMTFTYIGSSTWSPVQDNI-NETKIMH 408
 QY 485 TVNNTIPMSMCKRQSGQKPKVGIHVCCFECIDCLGTFLNHTEDEYECQACPNMWS 544
 DB 409 GKNOVPKSVSSDCLBEHQRYVTSFHHCCFECYCGAGTFLNKS-DLYKQPCCKEKA 467
 QY 545 YOSEISCFKROLVLELWHEAPTVALLAALGFLSTAILVLFMRHQPPIYRSAGPQC 604
 DB 468 PEGSOTCFPRVYFLALNEHTSWVLLANTLLLLLLITAGLFAHNDLPYVRSAGRLC 527
 QY 605 FLMLTLLVAYAVVYVYVGPVKVSTICRQALPFLCTICISCIASVSPQIVCAFKNAS 664
 DB 528 FLMLGSLAAGSGSLGFXEPFRPACILRQALFALGFIPLSLVRSDFLITIFKSTK 587
 QY 665 FPRAYSVWRQGPVYMAFIVLKAVIVIGMLARPOSHRTPDDBDKITIVGCNPYR 724

```

Db 588 VPEFYHAMVONHSGJLFVMISSAAQLICLTWLVMTPLPAREXORPPLHMECTET-- 645
QY 725 NSL-----LENTSLDLVSVGFSAFMGKELPTNYNEAKFTLSMTFEVFTSSVSLCTE 778
Db 646 NSLGFILAFIYXNG-----LBSISAFACSYLTKRDLPENTNEKCVFSLFFNFVSIATFTT 701
QY 779 MSAYSGVLTIVDLVTVNLALISLGYFGPKCYMLFPERMTPAYNSMIGCYTMR 836
Db 702 ASYVDKYLPAANMMAGLSLSGSGGYFLPKCYVILCRDLNSTEHPQASIDQYTRR 759

RESULT 11
073635
ID 073635 PRELIMINARY; PRT; 940 AA.
AC 073635;
DT 01-AUG-1998 (TEMBLrel. 07, Created)
DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE Calcium2+ sensing receptor.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98226788; Pubmed=9560249;
RA Naito T., Saito Y., Yamamoto J., Nozaki Y., Tomura K., Hazama M.,
RT "Putative pheromone receptors related to the Ca2+-sensing receptor in
RT fugu."
RL Proc. Natl. Acad. Sci. U.S.A. 95:5178-5181(1998).
DR EMBL: AB008857; BAA26122.1; -.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR000337; GPCR_Mgr.
DR InterPro: IPR000651; RasGEFN.
DR Pfam: PF00003; 7tm_3; 1.
DR Pfam: PF01094; ANF_receptor; 1.
DR PRINTS: PR00248; GPCRMRG.
DR PROSITE: PS00979; G_PROTEIN_RECPR_F3_1; 1.
DR PROSITE: PS00259; G_PROTEIN_RECPR_F3_4; 1.
KW Receptor.
SQ SQUONCE 940 AA; 105814 MW; 06DAB78036878B3 CRC64;

Query Match 23.9%; Score 1062.5; DB 13; Length 940;
Best Local Similarity 30.1%; Pred. No. 1.9e-79;
Matches 263; Conservative 172; Mismatches 336; Indels 103; Gaps 23;

QY 16 WTAEPAENSDFYLPDYLGLGLFSLHANKG-----IVHNFLOVPMCKEY 62
Db 16 YVISTYGPNOAOMTGILGLGFLPIHFGLISSKDNELANPESTKCVFNFRR----- 67
QY 63 EVKVIIGNLQAMRFAVEEINNDSSILPGVLGEIVDC-YISNNQOVLVPLA-HEDN 120
Db 68 -----GFRWQAMVFALEEINNSSSLPNTLGRITDCTNYSKALEATLSFPAQNKID 122
QY 121 LLLPIQE--DYSNTISRVAVIGPNSESVMVANFLSLFLPLQITYSAISDELNDKVRFP 178
Db 123 SLNDEDCNCTDHPATIAVGAAGSAVSTAVANLLSLFYTPQISVYASSRLLSNKQYK 182
QY 179 ALARTTSSAHVYAMQMLHFRMMWITLVSSDPTGRONGQLGRVARRDICIAFOE 238
Db 183 SFMKTIPTDHOATAMDAVLEYEQMNVIAVASDDDYGKRGIEKFEKEMERDICHINE 242
QY 239 TLPTLQPNQWTSBERQLTVIVDKLOQSTARVVVSPDLTYHFEFNEVLKRONFGAVW 298
Db 243 LISQY-----FDCETIALVDRIENSTAKYIVFASGPDLEPLIKEMVRRTDIRIW 294
QY 299 IASESNALDEVLNLTDELGLGTFTTISVDPGFSFERMGPOQAPPLSRTSQSYT 358
Db 295 IASEANASSSLIAKPEYLDVVEGIGVLAAGNIPGRFELQ---QVQPARGSINEVRE 351

```

```

QY 359 CNOECDNCL-----NATLSFNTILRLSGE-----RVVYSVYSAYV 393
Db 352 FWEETFCYLEDSPRIQESNGSDSPRLCTSEEDITSVEFPIIDHNLISVNYAVY 411
QY 394 AVAHALHSLIC-----DKSTCKRVVYPMOLLEIKVNFY-LLHQIFPDQGV 444
Db 412 SIHAALQDILICPPGHGLFANNSCADIKKEAQVQLKQLHLVYTNMGEKVFDEADM 471
QY 445 ALHEIYOWQMDSON--PROSVASYVPLQRO-----LKNIDISMHVNNVTIPMSCKR 498
Db 472 EANTITIMHRSABDGSVAFREVGYYHMAHARGAKLLIDNTKMMWNYSSVEPNSCED 531
QY 499 COSGOKKKPV-GIHVCCFECIDCLPGFTLNHTDEVECCQAPNNMWSYQSESCFQRLV 557
Db 532 CEPGRKGIIDSMPTCCFCECTECSDEYSDH-KDASICTCPNNSMSGHFTCFLEKE 590
QY 558 FLEHNEAPTAVALALAGLSTLAILVYWRHQPPIVYASAGPQCFMLTLILVAVNY 617
Db 591 FLAMSEPPGIALAICAVGLTAFVGVVFRFRNPPIVAKSNRELSTYLLSLICFSS 650
QY 618 VPIYVGPVKSTCLCQALPLCFTICISIAVRSPQIVCAFMAKSRFPRA-YSYWRYQ 676
Db 651 SLIFIGEPQWTCRLKQPAFGISFVLCISCLVKTNRVLVFE-AKIPISIRKMGLN 708
QY 677 GPVYSMAFIVLKNVIVIGML-ARQSHRTPDDPKRITIVSCNPYRNSLFTNTSLD 735
Db 709 LQELVFLCTFVOVMICVWMLYNAPSSYRNHIDE--IIFICNBSGVALFLLIGYTC 766
QY 736 LLSVGVFSFAYMCKELPTNYNEAKFTLSMTFEVFTSSVSLCTMSAV---SGVLVTVLD 792
Db 767 LLAICFFFAFKSKRLPENTENKFTFCMLIFFIYWI--FIPAFSTYGAFFVSAVEA 833
QY 793 LVTVLNLAISLG---YFGPKCYMLFPERMT 822
Db 824 IA-----ILASSYGLACIFPNKYIILFKPCRMV 853

RESULT 12
0923K1
ID 0923K1 PRELIMINARY; PRT; 858 AA.
AC 0923K1;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
DE Sweet taste receptor T1R3.
GN TAS1R3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH3T3;
RA Nelson G., Hoon M.A., Chandrasekar J., Zhang Y., Ryba N.J.P.,
RT "Mammalian Sweet Taste Receptors."
RL Cell 0:0-0(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH3T3; TISSUE=CIRCUVALLATE PAPILLA;
RX MEDLINE=21927605; Pubmed=11917125;
RL X., Staszewski L., Xu H., Dutler K., Zollner M., Adler E.;
RT "Human Receptors for Sweet and Umami Taste."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4692-4696(2002).
DR EMBL: AY032620; AAK51601.1; -.
DR EMBL: AF456324; AAM10636.1; -.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam: PF00003; 7tm_3; 1.
DR Pfam: PF01094; ANF_receptor; 1.
DR PROSITE: PS00980; G_PROTEIN_RECPR_F3_2; UNKNOWN_1.
DR PROSITE: PS00259; G_PROTEIN_RECPR_F3_4; 1.

```


Query Match	22.8%;	Score 1014;	DB 11;	Length 858;
Best Local Similarity	29.6%;	Pred. NO. 1.8e-75;		
Matches 250;	Conservative 164;	Mismatches 352;	Indels 80;	Gaps 22;

QY	24	NSDPEYLPEDYLLGGELFSTSHAMKRCITVHLEFNQV----	MCKEYEVKTVGYNULOMAMRPAVE	80
Db	26	SOQKRAQGDYTLGGELFPLGISTEET--LNOAORPNSTILCRF--	SPJGLFLAMAMKMAVE	81
QY	81	EINDDSLPLGVLLEYELVDCYISANNQVY-----LYFLAH-EDNULFLOEDYSNYIS	133	
QY	82	EINNSGALLPGLRIGYDLDFTC-----SEPVTWKSSIMFLAKVSGOSIAACNYQYOP	136	
QY	134	RVAVIYGDGNSQSVITVANEFLSTELLDQIYTSAISDELROKVRPALLRTTPSADHVEA	193	
Db	137	RYLAIVIGHSSELLLTGKFFSPFLMPOVYSASMDRISIRETFPSPFFRTPVSPSRVOLA	196	
QY	194	MYOLALHFRMMNIIYVSDPYGDNCOLLGERVARDICIAFOETLPTLOPNOMNMTSE	253	
Db	197	VYTLLOQNSWMVVAALSGDDYDGEGSLSPSSILANAGICIAHE----	GLYPOHDISGOQ	252
QY	254	RORLVTVYDKQOSTARVYVVEPDLTYHFEFNEVLNQNFTGAVIATASEMAIDPVLANL	313	
Db	253	LKATYDYLROVQNOBKQVQVYVLEAFARAVYSLFSTSIHHGLSPKVVWASEMLTSDLMWTL	312	
QY	314	TELGHLGTFELGIIIOVPIPESEFRMGPQAGRP-----LSTQSQSYCNQE	362	
Db	313	PNIARVGVIVLGFLORGALLPFEFSHYVETHIALADRPFCASLNAEJLDEHYMGORCP-	371	
QY	363	CDNCINATLSFNTILRLSGERV--VYSVYSAYVANAHALHSLLGCDKSTC-TRRVYPM	418	
Db	372	CDDIMLQNLMSGLLQNLNTSAGOLHHQIIPATYAAVYSVAQALHNTLOQVNSCHVSEHVLPM	431	
QY	419	QLLEIMKVNVTLLDHQIEPRPOGDVALHLEIYQOMODRQONPROQVASYRPLOROLKNI	478	
Db	432	QLLENNYKMSHADLTLQFPABENVDMEYDKMMVQOSPTPVLYHVTGTE----	NGTLQJL	487
QY	479	QDISMHTVYVNTIIPMSMCKSRQOSQOKRRPVGIVHCCECIDCLPGEFLNHTEDYEBOAC	538	
Db	488	QOSKMYMGQNVQPVQSSROCKDQGVBRVKGHSCCYDVCDAKAGSYRKH-PDDECTPC	546	
QY	539	PNNEMSVQSESCRKQOLVPLEMEHATIVALLAAGFSTALAIYIFRNHQTPYRS	598	
Db	547	NODQMSPEKSTACLPRRRKFLAMEBPVLSLLLLCLCVIGLIALAAGLSVHNHDSPLVOA	606	
QY	599	AGG-PMCFIATLTLVAVMYVVPVYVGPXYSTCICROALRPLCTTICISIAVRSQIYC	657	
Db	607	SGGSGOFCGGLCJCLFCLSTVL-LEPGRPSSASCLAQOPMAHLPLTIGLSLTLFLQAAETP-	664	
QY	658	AFKMAAREPRAYSTW---VRQGPVYSMAFTVULKVIVYIGMLARPQSHPRTPDP--	710	
Db	665	---VESELPJLSMANMLCSYLEGLMAMLYVLLATFEVALCAMLITAPP---PEVYDMSV	718	
QY	711	DPKTIIVSCNNYVNSLNFNTSIDLLLSVYVGSFSAVYKGLPPIVYNNAKITYLSMIFYFT	770	
Db	719	LPEYVLEHCHVRNSVNSLGLVHTNAMLAJCFGLTGFVLOSOPGRYNNARBLTPEAMLAFTI	778	
QY	771	SSVS-----LCFMSAY-----SGVLTVIIDLTVTVNLALASIGYGRPCYUMLIPEER	820	
Db	779	TWVSFVPLLANVQVAYQDPVAVQMGAILVICALGILVT-----FLLPRCYVLLMLPKL	828	
QY	821	NTPEAYF	826	
Db	829	NTQGEF	834	

```
Search completed: May 19, 2003, 09:52:37
Job time : 43.1811 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:48:40 ; Search time 14.3799 Seconds
(without alignments)
1714.646 Million cell updates/sec

Title: US-09-927-315-9
Perfect score: 4443
Sequence: 1 MGPRAKTICSLFILMLVLA.....ERNTPAYENSMIGYTRMRD 838

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3514	79.1	669	4	US-09-361-631-7 Sequence 7, Appli
2	3203.5	72.1	843	4	US-09-361-631-1 Sequence 1, Appli
3	3151.5	70.9	843	4	US-09-361-631-2 Sequence 2, Appli
4	1107	24.9	1059	4	US-09-134-513-2 Sequence 2, Appli
5	1094	24.6	1078	1	US-08-485-588-7 Sequence 7, Appli
6	1094	24.6	1078	1	US-08-485-588-7 Sequence 7, Appli
7	1094	24.6	1078	2	US-08-480-751-7 Sequence 7, Appli
8	1094	24.6	1078	2	US-08-943-986-7 Sequence 7, Appli
9	1094	24.6	1078	3	US-08-353-784-7 Sequence 7, Appli
10	1094	24.6	1078	3	US-08-484-719B-7 Sequence 7, Appli
11	1094	24.6	1078	4	US-08-484-159-7 Sequence 7, Appli
12	1092.5	24.6	1027	4	US-09-162-021B-2 Sequence 7, Appli
13	1088.5	24.5	1079	1	US-08-485-588-8 Sequence 8, Appli
14	1088.5	24.5	1079	1	US-08-484-565-8 Sequence 8, Appli
15	1088.5	24.5	1079	2	US-08-480-751-8 Sequence 8, Appli
16	1088.5	24.5	1079	2	US-08-943-986-8 Sequence 8, Appli
17	1088.5	24.5	1079	3	US-08-353-784-8 Sequence 8, Appli
18	1088.5	24.5	1079	3	US-08-484-719B-8 Sequence 8, Appli
19	1088.5	24.5	1079	4	US-08-484-159-8 Sequence 8, Appli
20	1087.5	24.5	1085	1	US-08-485-388-5 Sequence 8, Appli
21	1087.5	24.5	1085	2	US-08-484-565-5 Sequence 5, Appli
22	1087.5	24.5	1085	2	US-08-480-751-5 Sequence 5, Appli
23	1087.5	24.5	1085	2	US-08-943-986-5 Sequence 5, Appli
24	1087.5	24.5	1085	3	US-08-353-784-5 Sequence 5, Appli
25	1087.5	24.5	1085	3	US-08-484-719B-5 Sequence 5, Appli
26	1087.5	24.5	1085	4	US-08-484-159-5 Sequence 5, Appli
27	1079	24.3	1088	1	US-08-485-588-6 Sequence 6, Appli

28	1079	24.3	1088	1	US-08-484-565-6 Sequence 6, Appli
29	1079	24.3	1088	2	US-08-480-751-6 Sequence 6, Appli
30	1079	24.3	1088	2	US-08-943-986-6 Sequence 6, Appli
31	1079	24.3	1088	3	US-08-353-784-6 Sequence 6, Appli
32	1079	24.3	1088	3	US-08-484-719B-6 Sequence 6, Appli
33	1079	24.3	1088	4	US-08-484-159-6 Sequence 6, Appli
34	1012.5	22.8	877	4	US-09-619-353-2 Sequence 2, Appli
35	995	22.4	1219	2	US-08-687-289A-6 Sequence 6, Appli
36	917.5	20.7	863	4	US-09-619-353-14 Sequence 14, Appli
37	867.5	19.5	851	4	US-09-619-353-12 Sequence 12, Appli
38	856	19.3	854	4	US-09-619-353-10 Sequence 10, Appli
39	850	19.1	856	4	US-09-619-353-8 Sequence 8, Appli
40	849.5	19.1	835	4	US-09-619-353-7 Sequence 7, Appli
41	743	16.7	1058	2	US-08-687-289A-5 Sequence 5, Appli
42	724.5	16.3	915	1	US-08-453-862-2 Sequence 2, Appli
43	724.5	16.3	915	2	US-08-452-734A-2 Sequence 2, Appli
44	724.5	16.3	915	4	US-08-176-401B-2 Sequence 2, Appli
45	724.5	16.3	915	5	PCT-US94-14989-2 Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-09-361-631-7
; Sequence 7, Application US/09361631
; Patent No. 6383778
;
GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeyer, Juergen
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
; FILE REFERENCE: 02307E-08672005
; CURRENT APPLICATION NUMBER: US/09/361,631
; EARLIER FILING DATE: 1999-07-27
; EARLIER APPLICATION NUMBER: US 60/095,464
; EARLIER FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: US 60/112,747
; EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) B4 amino
; OTHER INFORMATION: acid sequence
US-09-361-631-7
Query Match 79.1%; Score 3514; DB 4; Length 669;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 666; Conservative 1; Mismatches 2; Indels 8; Gaps 2;
```

QY	162	ITVSAISDELARKVPPALFTTPSADHVEAMNOLMFRNNWIIYSSPTFYGRDNO	221
DB	1	ITVSAISDELARKVPPALFTTPSADHVEAMNOLMFRNNWIIYSSPTFYGRDNO	60
QY	222	ILGSEVARRDIDIAFOETPLTLOPNOMNTSEERORLVTVLKLQOSTARVVVFPDILL	281
DB	61	ILGSEVARRDIDIAFOETPLTLOPNOMNTSEERORLVTVLKLQOSTARVVVFPDILL	120
QY	282	YHFEVNEVLKONTGAVWVIAESWALDPVLAHNTLGLGLTGTITQSVPIPFGESEFRW	341
DB	121	YHFEVNEVLKONTGAVWVIAESWALDPVLAHNTLGLGLTGTITQSVPIPFGESEFRW	180
QY	342	GFOAAPPLSRTSOSYTCNOECDNCLNATLSNTILRLSGEVVYSVYSAAVVAHAHLS	401
DB	181	GFOAAPPLSRTSOSYTCNOECDNCLNATLSNTILRLSGEVVYSVYSAAVVAHAHLS	240
QY	402	ILGCDKSTCTKRVVYPMOLLEIRKVNFTLDDHQLFFDQGDVALHLIVQMWRSONP	461


```

? SEQ ID NO 2
? LENGTH: 843
? TYPE: prt
? ORGANISM: Mus sp.
? FEATURE:
? OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) B4 amino
? OS-09-361-631-2

```

Query Match	70.9%	Score 3151.5	DB 4	Length 843
Best Local Similarity	68.9%	Pred. No. 6.6e-291		
Matches 581	Conservative 113	Mismatches 142	Indels 7	Gaps 4

[illegible]

RESULT 4

US-09-134-513-2
Sequence 2, Application US/09134513
Patent No. 6210964
GENERAL INFORMATION:
APPLICANT: Brown, Edward M.
APPLICANT: Diaz, Ruben
APPLICANT: Bai, Wei
APPLICANT: Quinn, Stephen J.
TITLE OF INVENTION: The Avian Extracellular Calcium-Sensing
TITLE OF INVENTION: Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Vinson & Elkins L.L.P.
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20004-1008
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,513
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sanzo, Michael A.
REGISTRATION NUMBER: 36,912
REFERENCE/DOCKET NUMBER: BRI331/13003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)639-6585
TELEFAX: (202)639-6604
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1059 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHEICAL: NO
US-09-134-513-2

Query Match	24.9%	Score 1107	DB 4	Length 1059
Best Local Similarity	31.9%	Pred. No. 3.8e-96		
Matches 284	Conservative 148	Mismatches 354	Indels 104	Gaps 24

```

QY 9 CSEFLL--WVLAAPAESDPYLLGPDYLLGSEFLTHANMMGIYLNLOVPMCKEYEVK 66
Db 6 CCLILLEFYNWTAAYGPNQRAKKDIIIGSLFPIHF---GVAKD--QDLKSPEVEEC 60
QY 67 IGYNL-----MAMPAVEEINDDSLPGVLLGYELVDC-YISNNVQPLYFLA-HED 119
Db 61 IRYNFGFRWLOAMIFAEIEINNPNLLPNATLGYRFDTCNYSKSLXETLSLVAONKI 120
QY 120 NLBLEIQE--DYSNTISRVAAYICPDNSESVTVANFLSLELPQITXSATSELDREQVRE 177
Db 121 DSLNLDKDCENSEHILPTIANYVAGTSGSVTAVANLLGLFIIPQVSYASSRLLSNNOF 180
QY 178 PALIRTPSADHVAEAMVOLMLHRMMWIIIVLSSDTYGRDNGOLLGERVARDICIAFO 237
Db 181 KSPFRTIPNDEHQATAMADIIIEYFRMMVWGTAADDYGRGCIKREFEAEERIDCIDS 240
QY 238 ETLPLLOPNOMTSEERLOITVYDKLOOASTRAVVVESPDLTIYHFNEVLNRFNGAV 297
Db 241 ELI-----SQSDEBEELQVVEY---QNSARIAYVFFSSPDLEPLIKLEYVRNITGKI 292
QY 298 WIASSEMAIDVBLNHLTELGHGLTGITIOSVPIPGFSEFREMGPOAGPPPLSRFSQSY 357
Db 293 WIASEAWSSSLLIMPEFFRVIYGTIGTIFALKAGIOPFRETILQ---VYHFKRSANNFGAK 349

```

```

QY 358 TCNOCDNCL-----NATLSEN----- 374
: | |
: | |
Db 350 EMEETENCYLPSSEKNSPASAFHKAHEBGLGANGNGIAARPCOTGENTSVETPYMD 409
: | |
QY 375 -TILRSGERVVYSAVAVAHALSHLGC-----DKSTCKRVYPMOLLEETW 425
: | |
: | |
Db 410 FTHLRIS-----YVYLAIVASIAHALODIYCTPGKGLFTNGSCADIKKVAEMOVLKILR 464
: | |
QY 426 KVNFTL-LDHOIFPDPOGVALHLEIYOMOMDRSONP--FOSVASYPL-----OROLKNI 478
: | |
: | |
Db 465 HINTSMNGEYVDEFGDLGNTSIIMHLSPEDEGSVEEVEGHVNYAKKGERLNE 524
: | |
QY 479 ODISMHTVNNTPMSKSCRCQSGOKRPV-GIHVCCFECIDCLPGFLNTEDEYECOA 537
: | |
: | |
Db 525 NKILMSGSKVFPNSRDLPGTRKGIIEGPTCEGECVDCPDGEISDET-DASACK 583
: | |
QY 538 CPNNEMVQSOSTCFKQVLLENHHEAPTVALLAALGSLTALIVIFRHHQTPYR 597
: | |
: | |
Db 584 CPEDYMSNMENSCIPQIEFLSWTEPGIALTLFAVIGIFLTSFVLGVFKFRNTPYK 643
: | |
QY 598 SAGPMCFMLTLLVAVYVGPVKVSTCLROALPFLCPTICISIAVRSFOYC 657
: | |
: | |
Db 644 AFNHELSTLLFSLCCFSSSLFTEGEPQNTCHLRQPAFGISYVLCISCLIVKTRVYL 703
: | |
QY 658 AFKMASRPPRA-YSYWARYQGPVYSAFIVLKVIVIGMLARPOSHRTDPPDKITI 716
: | |
: | |
Db 704 VFE--AKIPTSLHRRMGLNQLFVLVLCFTVQIVICVIMVYTAAPSSYRNHELEDEIFE 761
: | |
QY 717 VSCPNRNSLLENTSLDILLISVGFSPAYMGKELPTNYNEAKITTSMTFFSVSLC 776
: | |
: | |
Db 762 ITCHGSLMAIGFLIGYCLLAICFFFAFKSRKLPENFNKAKRFTLSMLTFFIWMIS-- 819
: | |
QY 777 TFMASGCVLTIVDLTVYNLALISG---YEGPKCYMLFYPEENT 822
: | |
: | |
Db 820 -FIRAYASTYKQVY-AVEYAIILAAISGLACIFENKVIILKPKRNT 867
: | |

```

```

APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-588-7
Query Match 24.6%; Score 1094; DB 1; Length 1078;
Best Local Similarity 31.5%; Pred. No. 6,7e-95;
Matches 279; Conservative 160; Mismatches 342; Indels 106; Gaps 26;
QY 16 WTIAEPANSDYLP-----GDYLLGLFSLHANKGIYHNLQVPMCKEYEVKYG 68
: | |
: | |
Db 8 WVLLATWHTSAYGPDPAOKKGDIIIGLEPIHF--GVAARD--ODLKSRESVCEIR 62
: | |
QY 69 YNL-----MOAMRAVEEINNDSSLPLGVLGYEIVDVC-YISNNVQPVLYPLA-HEMDL 121
: | |
: | |
Db 63 YNERGFRLQAMTPAIEINSSPALNLTGLYIFDTCTVSAKALATTSFAQONKIDS 122
: | |
QY 122 LPIDQ--DYSNYISRYVAVIGPDNSESMTVANFLSLFLPOLITYSAISDELKDYRFP 179
: | |
: | |
Db 123 LNDEFCNCSBHIRSTAVVAGATSGSTAVANLGLFIYPOVYASSRLSKNQFKS 182
: | |
QY 180 LKRTPSADHVEAMVQMLHFRNMTIIVYSSPTVGRDNGQLGERVARDICIAQET 239
: | |
: | |
Db 183 FLRTIPNDEHQATMAIIIEFRNNWGTIAADDYGRPGIIEKREAEERDIDICIDFSEL 242
: | |
QY 240 LPILQPNQNTSEEROLATYIVDKLOOSTARVVVFPDGLTVLHFFNEVLRONFTGAVMI 299
: | |
: | |
Db 243 I-----SQYDEEITQVAVEYI--QNSTAKVIYVSSGPDLEPLKEIYRNITGKIWL 294
: | |
QY 300 ASESMAIDPYLHNLTEGLHGTGLTIGSVPIPGFSEF-----REMGPO 344
: | |
: | |
Db 295 ASEMASSSLIAMPQYHVAVGTTGFAKAGQJPGFREPLKTVAPRKSVANGFAKEWEE 354
: | |
QY 345 A-----GPPPLSTQOSTCNOECNCLNATISFTITILRSE----- 382
: | |
: | |
Db 355 TFNCHLOEGAKGPLVDTPLFRGH--EESGDRFSNSTAFRPL--CTGDENISSVEPYID 410
: | |
QY 383 -----RVVYSYSAVAVAHALSHLGC-----DKSTCKRVYPMOLLEETWKYNT 430
: | |
: | |
Db 411 YTHLRISTYNYLAYSTAHALODIYTLPGRLFTNGSCADIKKVAEMOVLKILRHNFT 470
: | |
QY 431 -LDHOIFPDPOGVALHLEIYOMOMDRSONP--FOSVASYPL-----OROLKNIODIS 482
: | |
: | |
Db 471 NNNGEYVTFPECGDLGVNYSIIMHLSPEDEGSYIFKEV--GYVNYAKKGRFLINEKIL 529
: | |
QY 483 WHYVNNTPMSKSCRCQSGOKRPV-GIHVCCFECIDCLPGFLNTEDEYECQACPN 541

```


QY 720 NPNYRNSLNTSLDLLSVGFSFAYMGKELPTNYNEAKFTLSMTFFTSVSLCTFM 779
 Db 766 HEGSIALMGLIGYTCILAAICFEFAFKSRKLPENFNEAKFTLSMTFFTSVSLCTFM 822
 QY 780 SAYSGVLTIVDLVYVNLNLAISLG----YFGKCYMILFYPERNT 822
 Db 823 PAVASTYGFKEVS-AVEVIAIILAAISFGILACIFFKKIYIILFKPSRNT 868

RESULT 7

US-08-480-751-7
 ; Sequence 7, Application US/08480751
 ; Patent No. 583684
 ; GENERAL INFORMATION:
 ; APPLICANT: Edward F. Nemeth
 ; APPLICANT: Edward M. Brown
 ; APPLICANT: Steven C. Hebert
 ; APPLICANT: Forrest H. Fuller
 ; APPLICANT: James E. Garrett, Jr.
 ; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: First Interstate World Center
 ; STREET: Suite 4700
 ; STREET: 633 West Fifth Street
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 90071
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: FASTSEQ
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/480,751
 ; FILING DATE: 7 June, 1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; PRIOR APPLICATION DATA: including application
 ; PRIOR APPLICATION DATA: described below: 9
 ; APPLICATION NUMBER: 08/353,784
 ; FILING DATE: 9 December, 1994
 ; APPLICATION NUMBER: PCY/US/94/12117
 ; FILING DATE: 21 October, 1994
 ; APPLICATION NUMBER: U.S. 08/292,827
 ; FILING DATE: 23 August, 1994
 ; APPLICATION NUMBER: U.S. 08/141,248
 ; FILING DATE: 22 October, 1993
 ; APPLICATION NUMBER: U.S. 08/009,389
 ; FILING DATE: 23 February, 1993
 ; APPLICATION NUMBER: U.S. 08/017,127
 ; FILING DATE: 12 February, 1993
 ; APPLICATION NUMBER: U.S. 07/934,161
 ; FILING DATE: 21 August, 1992
 ; APPLICATION NUMBER: U.S. 07/834,044
 ; FILING DATE: 11 February, 1992
 ; APPLICATION NUMBER: U.S. 07/749,451
 ; FILING DATE: 23 August, 1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hebert, Sheldon O.
 ; REGISTRATION NUMBER: 38,179
 ; REFERENCE/DOCKET NUMBER: 213/004
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (213) 489-1600
 ; TELEFAX: (213) 955-0440
 ; TELEX: 67-3510
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1078 amino acids
 ; TYPE: amino acid

; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-480-751-7

Query Match 24.6%; Score 1094; DB 2; Length 1078;
 Best Local Similarity 31.5%; Pred. No. 6.7e-95;
 Matches 279; Conservative 160; Mismatches 342; Indels 106; Gaps 26;

QY 16 WVLAEPAENSDPELP-----GDYLGLFSLHANKGIYVHNLQVPCKEVEYVIG 68
 Db 8 WVLALTWHTSAYGPDDRAOKKGDIIIGLFPPIHF---GVAARD--QDLKRSPEVCIR 62
 QY 69 YNL-----QAMRPAVEEINNDSSLPGLVEIYDVC-YISNNQPVLYFLA-HEDNL 121
 Db 63 YNFRGFMLOAMIPAIIEINSSPALPLNLLGRIPTCVTSKALBATLSFVQNRKIDS 122
 QY 122 LPIDF--DYSNYISRYAVVIGPDSESVMYANLSFLPLQITYSAISELNDKVFPA 179
 Db 123 LNLDFCNCESEHLPSTIAVVGATGSGVSTAVANLLGLFYIPOVSASSRLLSKRNFKS 182
 QY 180 LLRTTPADHHEAMVOLMLHFRNMTIIVSSPTYGRDNGQLGERRVARDICIAFOET 239
 Db 183 FLRTIPDEHQATMADITIEFRNMTIIVSSPTYGRDNGQLGERRVARDICIDISEL 242
 QY 240 LPTLQPNQNTSEERQRLVTVDKLQOSTARVVVSPDPLTLHFNEVLRFNFTGAWTI 299
 Db 243 I-----SOYSEDEIHQVVEYI---QNSTAKVIYVSSGPDLPLKEIYRRNTGKIWL 294
 QY 300 ASESMALDPVNLNTELGHTGFIQTISVPIRGESEF-----REMGPO 344
 Db 295 ASEMASSSLIAMPQYFHVVGTTGFALKAQOIPGFEPLKRYAPKRSVHNGFAKEWEE 354
 QY 345 A-----GPPPLRTSOSYTCNOECNCLNATISFTYILRSGE----- 382
 Db 355 TFNCHLOEAGKPLPVPTFLRGH--EESGDRFSSTAFAPRL--CTQDENISSVEPYID 410
 QY 383 ----RVVYSYSAVYVAHAHLSLGC-----DKSTCKRVVYVWOLLEIMKYNFT 430
 Db 411 YTHLRISYNYLAVYSIAHLODIYTCPLGRGLTFPNSCADIKKVEAMOVYLKHLRHNFT 470
 QY 431 -LIDHOLFPPQGVNLAHLEIYQWMDRSONP--FQSVASYPL-----OROLKNIDIS 482
 Db 471 NNMGEOYTFDECGDLVGNYSITLNNHLSPEDEGSIFYKEV--GYINYAKKGERLFTNEKIL 529
 QY 483 WHYVNTTIPMSMCKRSQSGOKKRPV-GIHYCEPCIDCLPGFLNHTEDYEEOACPPNN 541
 Db 530 WSGSRREVPVSNCRDCLAGRKGIIEGPYCECEVCEPDGEYSDET--DASACNKPDD 588
 QY 542 EWSIQSETSCPKROLVLEHNEAFTIYVALLAALGFLSTALIVTFRKHOTPIVRSAG 601
 Db 589 FWSNENHTSCIAKIEIEFLSWTEPGIALTFEAVIGIFLTFVGLVFTKFRNTPIVKATNR 648
 QY 602 PNCPLMTLLLVAAVYVAVVGPVKVSTCROALPFLCTICSCIAVRSFOVCAF-- 659
 Db 649 ELSTLLFLSLCCPSSSLFTIGEPQDWTCLRKQAFISIVYLISCILVATNRLLVFEA 708
 QY 660 KMSRFPRAVSYWRYOGPYVSMATITVLMKVIYVIGMLARPOSHPRTPDDDPKITIVSC 719
 Db 709 KIPISFHRK---WNGMLQGLVFLVFCFMQIVICVIMLYIARPSYNOELEDIEDIIFITC 765
 QY 720 NPNYRNSLNTSLDLLSVGFSFAYMGKELPTNYNEAKFTLSMTFFTSVSLCTFM 779
 Db 766 HEGSIALMGLIGYTCILAAICFEFAFKSRKLPENFNEAKFTLSMTFFTSVSLCTFM 822
 QY 780 SAYSGVLTIVDLVYVNLNLAISLG----YFGKCYMILFYPERNT 822
 Db 823 PAVASTYGFKEVS-AVEVIAIILAAISFGILACIFFKKIYIILFKPSRNT 868

RESULT 8

US-08-943-986-7
 ; Sequence 7, Application US/08943986
 ; Patent No. 5962314

```

GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,986
FILING DATE: 03-OCT-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/484,565
FILING DATE: 7-June-1995
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-986-7

Query Match 24.6% Score 1094 DB 2 Length 1078
Best Local Similarity 31.5% Pred. No. 6.7e-95
Matches 279; Conservative 160; Mismatches 342; Indels 106; Gaps 26;

```

```

122 LPIQ--DYSNYSRVAIVIGPDNSESVMTVANFLSLFLPQITYSATSDLRKYRPA 179
123 LNLDFCNCSEHPTPIAVVATGSGVSTAVANLGLFTIPVSTASSRLSNQFNS 182
140 LIRTPSADHVEAVQMLHFRMWTIVLVSSDTYGRDNGOLLERVARDCIAFOET 239
183 FLRTIPNDHQATAMADILEYFRMNVGTAADDDYGRGIKFRFEAEERDIDFSEL 242
240 LPTLPQNMNSEERQRLTYDKLQSGARVYVPSDPLTYHFNFLRQNFQAWI 299
243 I-----SOYSDEEIOHVEVI---ONSTAKYIVVFSSGPDLEPLIKETVRNITGKTL 294
300 ASESNAIPVLNHLNTELGITGTTIOSYVIPFSEF-----REMGPO 344
295 ASENANSSSLIMPOYFHVVGSTIFALKAGIIPGFRFLKRVHNRKSVHNFAPAEFWE 354
345 A-----GPPLSRITSGSYTCNQECNCLNATLSEFTLLRSGE----- 382
355 TFNCHLQEGAKGPLVPDFLRGH--EESGDRFSNSTARPL--CTGDENISSVETPYID 410
383 ----KVISYSAVAVAHALSLGC-----DKSCTKRVYYPQMLEIKVNF 430
411 YTHLRISYNVYLAYSIAHALODIYTCPLGRGLFTNGSCADIKYEAQVLAHLRLNFT 470
431 -LLHQIFPDPOGDVALHLEIYQWOMDRSQND--FOSVASYPL-----ORQKNIODIS 482
471 NMGEOVTFDECGDLVGNSTIINHLSPEDGSIVREV--GYNVYAKKGERLFINDEKTL 529
483 WHTVNTIPMSCKRCSGQARKRPV-GIHWCCFPCIDCLPCTPLNHEDEVECOACPN 541
530 WSGFREVPEVNSCRDCLAGTRKGIIEGPCTCFECPCDEPDEYSDET--DASACNKPDD 588
542 EMSYQSEISCFKRLVFLFMHAPITIAVALLAFLSLTALIVFMKRFOTPIYRSAG 601
589 FMSNHNHTSCIAKEIEFLSWTEPGIALTLFVLDIIFLAFLGVEIFRNPPIYKATNR 648
602 PNCFLMLTLIAVAVVPPVYGPVKSICIQALFPLCTFCICIAVRSQIACAF-- 659
649 ELSTYLLPSLCCFSSSEFIEEPDWTCLRLQPAFISFVICISICILKTRVLLVPEA 708
709 KIPTSFHRK---WGLNLQFLFLVFLCTEQIYICVIMLYTAPSSYRNQELDEILFIC 765
720 NNTYNSLLENTSLDLSVGFSPAYMKELEPTYNNAKFTTISMTFTTSSVSLCTM 779
766 HEGSLMALGELIGYTCCLAALICFFFAFSRKLPEMNEKRTFTFSMLIFFIWIS--FI 822
780 SAYSGVLVTVIDLTVNLILALISLG---YFGPKCYMILFPERNT 822
823 PAVASTYCKFVS-AVEVAIIILASFGLLACIFFNKIIILFRPSNT 868

RESULT 9
US-08-353-784-7
Sequence 7, Application US/08353784
Patent No. 6011068
GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth, Edward M.
APPLICANT: Brown, Steven C. Hebert,
APPLICANT: Bradford C. Van Wageningen, Manuel
APPLICANT: F. Balandrin, Forrest H. Fuller,
APPLICANT: Eric G. Delmar, and Scott T. Moe
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles

```

```

STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,784
FILING DATE: 9 December, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 8
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 209/069
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-784-7

Query Match      24.6%; Score 1094; DB 3; Length 1078;
Best Local Similarity 31.5%; Pred. No. 6,7e-95;
Matches 279; Conservative 160; Mismatches 342; Indels 106; Gaps 26;

QY 16 WVLAEPAENSDFYLP-----GDYLLGGLFSLHANMGIYHLPVPMCKEYEVKIG 68
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 8 WVLALMTWHTSATYGPDDQRAQKKDIIIGLFPPIHF---GVAAKD--QDLKRPSVEGIR 62
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 69 YNL-----MQAMRAVEAEINNDSSILPQVLLGYEIVDVC-YISNNVQVLYELA-HEDNL 121
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 63 YNFGFRLAQMIPIAEIRINSPPALPMLTLGYRIFDCTNVSXKALETSLFVAQNKIDS 122
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 122 LPIDQ--DYSNYISRYAAVIGPDNSESEMYANFLSLFLQLQIITYSAISDELKQVRPA 179
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 123 LNLDFECNCSHIPSTIAVAGATSSGVTAVALNLIQLEIYIPQVSVASSRLSNKQKRS 182
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 180 LRLTPSADHVEAMVOLMLFRNNMIIIVVSSDYGNDGQLLGERVARBDICIAFOET 239
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 183 FLRLIPNDEHQAATMAADIEFRKNWGTIAADDDYGRGIEKREKEEDIDICIDFSEL 242
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 240 LPTIQPNOMNTSEERQRLVTIVDKLOSTARVAVVSPDLTLVYHFNELRQNTGAVWI 299
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 243 I-----SOYSDDEEIQHVEVI---QNSTAKIYVSSSGPDLEPLIKETIVRNITGKIML 294
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
QY 300 ASESMADIPVHNLTELGLTFLGITIGSVPIPGFSF-----REMGPO 344
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

```

Db 295 ASEMASSSLIAMPQYFHVVGITIGFALKAGQIPGFEFLKRVPRKSVHNGFAKEWEE 354
QY 345 A-----GPPPLRTSOSYCNQEOCNCLNATLSFNTILSGE----- 382
Db 355 TFNCHLOBGAKPPLPVOTFLRGH--EESGDRFSNBSIAFRL--CTDENINSVETPYID 410
QY 383 ---RVYSVYSAYVAVAHALSHLLG-----DKSTCKRVYVPMOLLEIKVNYFT 430
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 411 YTHLRISVNYLAVYSTAHALODIYTCIPRGFLTNSCADIKVEAMOVYLKHLRHLNFT 470
QY 431 -LIDHQIFPDQGVVALHLEIVQWQMDRSQNP--FQSYASTYPL-----QRLKNIDIS 482
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 471 NMGEQVTFDECQDGLVNTSTIINHLSPEDGSIYKEV-gYVNYARKGEFLINEKIL 529
QY 483 WHYVNTIPLSMSCSKROSGOKRPV-GIHWCCPECIDCLPGFTLNHTEDEYECQACPNN 541
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 530 WSGFREVPSNCSRDCLAGRKIIIGEPICCEVCYCPDGEISDET--DASACNKPDD 588
QY 542 EMSYQSETSCFKQVLEWHHEAPTLVALLAAGFLSTAILVIFWRHQTPIVRSAGG 601
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 589 FWSNENHTSCIAKIEIEIEMTEPEGIALTLFAVLGIFLAVLGFKFRNTPIVKATNR 648
QY 602 PMCFMLTLILVAMVVPVYVGPVKVSTCLCRQALFPLCFITCICIAVNSFOIVCAF-- 659
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 649 ELSTLLFSLCCFSSSLFEGEPQDWTCLRQPAFISFVLICSLILVKNRVLVFEA 708
QY 660 KMASFPRAYSVWVRYOGPYVSMAFITVLMKVIYIGMLARPOSHPTDPPDKRTIVSC 719
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 709 KIIPSFRK---WGLNLQFLVFLCTFMQIVOLCVIMLYTAPPSYNOLEEDILITTC 765
QY 720 NPNYNSLLENTSLDILLISVYGFSPAYMGKELPNYNEAKFTYLSMPTFTSSVSLCTFM 779
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 766 HEGSMLALGFLIGYTCILAAICFEFFAFKSRKLPENFENAKFTYFSMILFTVWIS--FI 822
QY 780 SAYSGVLTVIDLVTVNLILALISLG---YFGKCYMILFYPKRT 822
   |||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 823 PAIATYGRKTVS-AVEVIAIHAASFGILLACITFNKITYIILFKPSRNT 868

RESULT 10
US-08-484-719B-7
Sequence 7, Application US/08484719B
Patent No. 6031003
GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth, Edward M.
APPLICANT: Brown, Steven C. Hebert,
APPLICANT: Bradford C. Van Wageningen,
APPLICANT: Manuel F. Balandrin,
APPLICANT: Forrest H. Fuller, Eric G.
TITLE OF INVENTION: CALCULUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
City: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS Word
SOFTWARE: FASTSEQ for Windows Version 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,719B
FILING DATE: 7 June, 1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/353,784

```

FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Douglas C. Murdock
REGISTRATION NUMBER: 37,549
REFERENCE/DOCKET NUMBER: 213/007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-7198-7

Query Match 24.6%; Score 1094; DB 3; Length 1078;
Best Local Similarity 31.5%; Pred. No. 6.7e-95;
Matches 279; Conservative 160; Mismatches 342; Indels 106; Gaps 26;

QY 16 WTLAEPANSDYLP-----GDYLLGLGFLSHNMKGIVLNFLOVPMCKEYKVG 68
DB 8 WLLALHTHTSAVGPDRQKKGDIILGGLFIHF---GVAANKD-QDIKSPRESVECHR 62
QY 69 YNL-----MQAMRAVEEIEINDSLLPGVLLGEYIVDC-YISNNVQVLYELA-HEENL 121
DB 63 YNRFGRMLQAMIFAEIEINSPLLPNTLGLRIFDTCNTVSKALEATLSVAQNKIDS 122
QY 122 LPIDQ--DYSNTISRVAVTGPDNSESVTVANFLSLPDIPTTSAIDELDKYRPPA 179
DB 123 INLDEFCNCSEHIFSTIAVAGTGSVSTAVANLLGLFYIPQVSYASSRLLSNKQFES 182
QY 180 LRTTPSDHVEAMVQLLHFRMMWIIYLVSSDTYGRDNGOLLGERVARRDICIAFORT 239
DB 183 FLRTIPNDHQATAMADIEYFRMNVGIIAADDDYGRGIEKFRFEAEERICIDFSEL 242
QY 240 LPTLPNQMNTSEKQRLVTIVDKLQOSTARVVVFPDLTLTHFENEVLKONFTGAWVI 299
DB 243 I-----SQYSDDEEIQHYEVI---QNSTAKYIVVSSGPDLEPLIKEIVRRNIGKIML 294
QY 300 ASESNAIDPVNLNLELGLTGLTITQSVIPIGFSER-----KEMGQ 344
DB 295 ASEAWASSLLMPQIFHVHGGTIGPALKAGQIPGRRELLKRVHPKRSVHNGFAEFWEE 354
QY 345 A-----GPPLSRTSQSYTCNOECDNCLNATLSFNTILRLSGE----- 382
DB 355 TENCHEGAKPLPYDFLRLGH--ESGDRFSNNSSTARPL--CTGDNISSVETPYID 410
QY 383 ----RVYSVYSAYVAVAHALSLGCG-----DKSTCTKRVVYPMQOLEEIMKVNFT 430
DB 411 YTHLRISYVYLVAVYSIAHALODIYTCLPRGLFTNGSCADIKKYEAMQVLRHLHLNFT 470
QY 431 -LLDHOIFEDPDQGVALHLEIYQWQMDRSONP--FOSVASYPL-----QRLKNIQDIS 482
DB 471 NMGEOVYTFDECGDLVGNSTITNMHLSPPDGSIIVREV--GYINVYAKKGERLPIKEKITL 529

QY 483 WHTVNTTIPMSCKRCQSGOKKKPV-GIHVCCFECIDCLPGLFNLHTEDEYEQACPRNN 541
DB 530 WSGFNEPEVFNCSNRCLAGTRKGIIEGEPTCCFCEVECPDEYSDET--DASACNKCPRD 588
QY 542 EMSYQSESCFKROLVFLFMHEAPITAVALLAAGFLSTLALIVFWRHFOPIYRSAGG 601
DB 589 FSNENHISCIKAEIEFLSWTEPFGALTLFPAVLGIFLTAFLVAGYIKRKNPIYKATNR 648
QY 602 PRCFLMLTLLVAVNVPYVGPVKVSTCLQALPELCTTICISCIQAVRSQIVCAF-- 659
DB 649 ELSYLLLEFLCCFSSLEFFIGEPQWTCRLQPAFGISFVICISCIITKTRVLLVFEBA 708
QY 660 KMASRPRATSYWVRVQGGYVSMATFVTKMTIVYIGMLARQSHRDPDDPKTIYVSC 719
DB 709 KIPTSFHRK---WGLNLQGLFLVFLCTFMQIVCIWLTAPSSYRNQDELEDEITFTC 765
QY 720 NPNYRSLFNTSLDLSVGFSPAYMGEKLPYVNEAKFTTLSMTFFTSVSLCTPM 779
DB 766 HEGSLMALGFLGYTCULALICFFPAFKRKLPENNEAKFTTFSLIFFIYWIS---PI 822
QY 780 SAYSGVLTIVDLTVNLALISLG---YGPCKYMLFYPERNT 822
DB 823 PAYASTYGRFVS-AVEVIAIILASFGLLACIFPNKIYIILLFRPSRNT 868

RESULT 11

US-08-484-159-7
Sequence 7, Application US/08484159
Patent No. 6313146

GENERAL INFORMATION:

APPLICANT: Bradford C. Van Wageningen
APPLICANT: Manuel F. Balandrin
APPLICANT: Eric G. Del Mar
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,159
FILING DATE: 7 June, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9

APPLICATION NUMBER: 08/353,784

FILING DATE: 9 December, 1994

APPLICATION NUMBER: PCT/US/94/12117

FILING DATE: 21 October, 1994

APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994

APPLICATION NUMBER: U.S. 08/141,248

FILING DATE: 22 October, 1993

APPLICATION NUMBER: U.S. 08/009,389

FILING DATE: 23 February, 1993

APPLICATION NUMBER: U.S. 08/017,127

FILING DATE: 12 February, 1993

APPLICATION NUMBER: U.S. 07/934,161

FILING DATE: 21 August, 1992

APPLICATION NUMBER: U.S. 07/834,044
 FILING DATE: 11 February, 1992
 APPLICATION NUMBER: U.S. 07/749,451
 FILING DATE: 23 August, 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Hebert, Sheldon O.
 REGISTRATION NUMBER: 38,179
 REFERENCE/DOCKET NUMBER: 214/101
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1078 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-484-159-7

Query Match 24.6%; Score 1094; DB 4; Length 1078;
 Best Local Similarity 31.5%; Pred. No. 6,7e-95;
 Matches 279; Conservative 160; Mismatches 342; Indels 106; Gaps 26;

16 WTAEAPENDSYLP-----GDYLLGLFSLHANKGIYHNLQYPMCKEYEVYIG 68
 8 WLLATLWHTSAYGPDRQAKKGDIIIGLFPIHF---GVAKD--DOLKSRPESVEICR 62
 69 YNL-----MOMRFAVEINNDSSLRVLLGEIVDYC-YISNNQPVLYFLA-HEHNL 121
 63 YNFGFRLQAMIFAIIEINSSPALLPLTLGYRIFDCTNVSALERTLSFVANKIDS 122
 122 LPIDQ--DYSNYISRYVAIVIGPDNSESVMYANFLSLFLPQIYSAISDELARKVREPA 179
 123 LNDEFCNCSHISTAVATGSGSVTAVANLGLFYIYQVSYASSRLSKNKQPKS 182
 180 LKRTPSADHVEAMVOLMHRMNIIVLVSSDTYGRDNGQLGERVARDICIAFOET 239
 183 FLRTIPNDEHQTAMADIEYFRMNVGTIAADDYGRPIEKRFEAEEDICIDFSEL 242
 240 LPTQPNOMTSEERORLYTVDKLOOSTARVVYVSPDLTLHFNFVNLQNTGAVWI 299
 243 I-----SOYSDDEIQAHEVY---QNSTAKVIVFSSGPDLEPLIKEIVRNITGKWL 294
 300 ASESMAIDPVLANLTELGLTFLGITIOSVPIPGFSEF-----REMGPO 344
 295 ASEMASSSLAMPQYFHVVGITGIPALKAGQIPGRFLKVKPRKSVHNGFAKEFWE 354
 345 A-----GPPPLRTSQSYTCNOBDCNCLNATLSFNTILRISGE----- 382
 355 TFNCHLOEGAGPLPVDFTFLRGH--EESGDRFSNSSTAFRPL--CTGDENISSVETPYID 410
 383 -----RVVYSYSAVYAAVAHLSLGC-----DKSTCKRKYVPMQLLEIMKAVFT 430
 411 YTHLRISTNYLAVYSIAHALQDIYTCIPGRGLFTNSCADIKKVEAMQVILKRLHNFT 470
 431 -LIDHQTPEQDVALAHLIELVQMDRSONP--FOASVASYPL-----OHOLKNIDIS 482
 471 NNMEQYTFDECDGLVGNYSIINMHLSPEDGSYFKEV-GYNNYAKKGEELFINEKITL 529
 483 WHYNNNTIPMSCKSKRCQSGOKKRPV-GIHVCECICIDCLPGTFLNHTDEYECQACPNN 541
 530 WSGFSREVPFSCSRDCLAGTRKGIIEGEPICCEVECPGEYSDET-DASACNKCPDD 588
 542 EWSQSTSECKRQVLFLMEHETIAVALLAAGFLSTALIVYFMRHPTPIYRSAGG 601
 589 FWSMENTSCIAKEIEELSWTEPFGIALTLFAVIGIFLTAFLVGLFKFRNTPIYKATNR 648
 602 PMCIAMTLLILVAVVYVYGPVKYSTCLCRQALPFLCICISCAVRSFOVCAF-- 659
 649 ELAVLLFLSLCCFSSSLFTFGEQDWTGRLRQPAFGISFLYCLISCLIVKKNRLVFEA 708
 660 KMASRFPRAVSYYWYOGPYVSMAFITVLKNIIVVIGMLARPOSHPRTPDDPKITIVSC 719

709 KIPIFSFRK---WMGLNIQFLVFLCFTFMQIVICVIMLYTPAPSSYNOELEDEIRITFC 765
 720 NPNRNSLFTSILDLISVYGFSAFMKRELPNNYNEARFITLSMFFPSSISLCTFM 779
 766 HEGSIMLGLIGYTCILAAACFFFAKSRRLPNNFNEARFITSMLIFLWIS--FI 822
 780 SAYSGVLVTVYIDLIVYVNLALISLG---YFGRCMILFEPRENT 822
 823 PAVASTYGRKVS-AVEYVAILAASFGLLACIFENKIIITILFKPSRNT 868

RESULT 12

US-09-162-021B-2

Sequence 2, Application US/09162021B

Patent No. 6337391

GENERAL INFORMATION:

APPLICANT: H. William Harris

APPLICANT: Edward M. Brown

APPLICANT: Steven C. Hebert

TITLE OF INVENTION: Polycation-Sensing Receptor in Aquatic

TITLE OF INVENTION: Species and Methods of Use Thereof

FILE REFERENCE: 2856,1001-007

CURRENT APPLICATION NUMBER: US/09/162,021B

CURRENT FILING DATE: 1998-09-28

PRIOR APPLICATION NUMBER: PCT/US97/05031

PRIOR FILING DATE: 1997-03-27

PRIOR APPLICATION NUMBER: 08/622,738

NUMBER OF SEQ ID NOS: 19

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 1027

TYPE: PRT

ORGANISM: squallus acanthias

US-09-162-021B-2

Query Match 24.6%; Score 1092.5; DB 4; Length 1027;
 Best Local Similarity 31.0%; Pred. No. 8.6e-95;
 Matches 277; Conservative 162; Mismatches 360; Indels 95; Gaps 25;

5 AKTICISFELIMVTAEPENDSYLP-----GDYLLGLFSLHANK--KGIYHNLQ 55
 2 AQLHCQLFLFETLQ--SYNVSIGYPNORAKKGDIIIGLFPIHFVGAKKDOLKSRPE 60
 56 VPMCKEYEVKIGYVNLQAMRFAVEEINNDSSLRVLLGEIVDYC-YISNNQPVLYE 114
 61 ATKCIIRYNER--GRWIDQAMIFAIIEINSSKTEPLNITLGRIDTCNTVSKALEATLSF 118
 115 LA-HEHNLPIQD--DYSNYISRYVAIVIGPDNSESVMYANFLSLFLPQIYSAISDEL 171
 119 VAQNKIDSLNDEFCNCSHISTAVATGSGSVTAVANLGLFYIYQVSYASSRL 178
 172 RDKYRFPALLRTTPSADHVEAMVOLMHRMNIIVLVSSDTYGRDNGQLGERVARD 231
 179 SNKNEYKAFLETPNDQOATMAEIEHFQOMNVVGTIAADDYGRGIDKFRREAVKRD 238
 232 ICIAFOETLQPNOMTSEERORLYTVDKLOOSTARVVYVSPDLTLHFNFVNLQ 291
 239 ICIDFSEMT-----SOYTT--QKQLEFIADVIONSSAKVIVFSSGPDLEPLIQLTVR 290
 292 NFGVAVIASESMALIDVLANLTELGLTFLGITIOSVPIPGFSEF----- 340
 291 NIDIRIWLASEMASSSLINKPEFHVVGITIGFALAGRIPGNKRKLKVEHPSRSSDNG 350
 341 -----W-----GPOA-GPPPLRTSQSYTCNOBDCNCLNAT 370
 351 FVKEFEWETPNCYTEKTELTOLKNSKVPSPAAOGSGSKAGNSRRRLAHPHPCGEENIT 410
 371 LSFNTILRISGERVYVYSAVYAAVAHLSLIGCDST-----CTK--RVYVPMQLE 422
 411 SVETPYLDYTHLRISTNYVAVYSIAHALQDIHSCPKGTSIFANGSCADIKKVEAMQVLN 470

Db 642 IYKATNELSYLLFLSLCFSSSLFIEGPDQWTRCLROPAGISFVLCISCLIVYTNR 701
 QY 655 IYCAF--KMASRPRASVWRYGQPVYSAFIVLKWIVYIGMLARPOSHPTDDDP 712
 Db 702 VLVFEAKIPTSFHRK--WMLGLOFLVFLCTFMOILCTIWLTAAPSSYNHHELD 758
 QY 713 KITVSCNPVYRNSLFTNTSL--DLLSVYGFSAFANGKELPYNVEAKFITLSMTFYF 769
 Db 759 EIIPTTC--HEGSLMALGSLIGYTCLLAICFPFAKRSRLPENFNAKFTFSMLIF 815
 QY 770 TTSVSLCTFMSAYSGVLTVYIDLTVNLALISLG---YFGKCYMILFYPERNT 822
 Db 816 IYWIS--FLPVAYSTYGRFVS-AVEYIALAASFGLLACIFENKYYIILFKPSRNT 868
 RESULT 14
 US-08-484-565-8
 Sequence 8, Application US/08484565
 Patent No. 5763569
 GENERAL INFORMATION:
 APPLICANT: Edward M. Brown
 APPLICANT: Steven C. Hebert
 APPLICANT: James E. Garrett, Jr.
 TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 TITLE OF INVENTION: MOLECULES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: First Interstate World Center
 STREET: Suite 4700
 STREET: 633 West Fifth Street
 CITY: Los Angeles
 STATE: California
 COUNTRY: USA
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: FASTSEQ
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,565
 FILING DATE: 7 June, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 PRIOR APPLICATION DATA: including application
 PRIOR APPLICATION DATA: described below: 9
 APPLICATION NUMBER: 08/353,784
 FILING DATE: 9 December, 1994
 APPLICATION NUMBER: PCT/US/94/12117
 FILING DATE: 21 October, 1994
 APPLICATION NUMBER: U.S. 08/292,827
 FILING DATE: 23 August, 1994
 APPLICATION NUMBER: U.S. 08/141,248
 FILING DATE: 22 October, 1993
 APPLICATION NUMBER: U.S. 08/009,389
 FILING DATE: 23 February, 1993
 APPLICATION NUMBER: U.S. 08/017,127
 FILING DATE: 12 February, 1993
 APPLICATION NUMBER: U.S. 07/934,161
 FILING DATE: 21 August, 1992
 APPLICATION NUMBER: U.S. 07/834,044
 FILING DATE: 11 February, 1992
 APPLICATION NUMBER: U.S. 07/749,451
 FILING DATE: 23 August, 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Hebert, Sheldon O.
 REGISTRATION NUMBER: 38,179
 REFERENCE/DOCKET NUMBER: 213/006
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1079 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-484-565-8
 Query Match 24.5%; Score 1088.5; DB 1; Length 1079;
 Best Local Similarity 31.7%; Pred. No. 2.2e-94;
 Matches 284; Conservative 164; Mismatches 328; Indels 121; Gaps 31;
 Db 9 ALALAN-----HSSAYGPDQRAKKGDIIIGGLEPIH---GVAKD--QDLKSRPE 56
 QY 63 EVKVIYNL-----MQAMRAVEEINNDSSILPGVLLGLEYIVDC-YISNVQPVLYFLA 116
 Db 57 SVECIKRYNFRGFWLQAMIAIEINSSPSILPMWTGRIYFDICNTVSALNTLSFVA 116
 QY 117 -HEDNLPIDQ--DYSNYSRVAVVAVIGPDNSESVMVAFNLSFLPQITYSATSDLRD 173
 Db 117 QNKIDSLNDEFCSCSHIPSTIVAGTSGVSTANALGLFYITQVYASSRLSN 176
 QY 174 KVRPALLRTTPSADHVEAVQMLHFRNMWITIVSSPTGRCNQLGERRYARDIC 233
 Db 177 KQYKSFRLTIPNDHQATMAADIEFFRMWVGTIAADDYGRGIEKFEKEAEERDIC 236
 QY 234 IAPQETLPTLOPNOMTSEORLTVDKLOOSTARVVVPSDDLHYHFNFLVLRQNF 293
 Db 237 IDESELT-----SOYSDEELQOVVEVI--QNSTAVIVVFSSGPDLEPLIKETVRNI 288
 QY 294 TGAVWIASGWAIDPVNLNTELGH-LGTEFGITIQSVPIPGSEF----- 338
 Db 289 TGRITWLSSEAMASSSL-ANPEYHVAVGRTIGFGLKAGQIPGFEFQKXHPKSVANGF 347
 QY 339 -REWGPOA-----GPPUSRTQSSTYCNQECNCLNATLSFNTILRSGE----- 382
 Db 348 AKEFMEETFNCHLOEGAKGLPVDTFVRSH--EEGGRNLNLSVAFRPL--CTGDEINIS 403
 QY 383 -----RVVYSVYSAVYAVAHALHSLGC-----DKSTCKRVVYPMOLLEE 423
 Db 404 VEPYMDYEHIRLISYNYLAVYSTAHALODITYTLPRGLFTNSCADIKVEAMOVLYKH 463
 QY 424 IWKVNT-LDHOJFDPQGDVALHLEIVOMWDRSONP--FOSVASYPL-----OROL 475
 Db 464 LRHLNFTNMGEQVTFDECGDLVGNYSILIMWHLSPEDGSIYKEV--GYUNVYAKKGERLF 522
 QY 476 KNIDDSMHVYNNITIPMSCKRSQSQKKPV-GIHYCCPECIDCLPGFLNTEDEYE 534
 Db 523 INEKILMSGFSREVPSNCRDQAGTRKGIIEGEPTCCFCEVCEPDGEYSGET--DASA 581
 QY 535 COACPNNMWSYQSESCFRQLVLEHWEAPTIYVALAALGFLSTALIVYFHRHPTP 594
 Db 582 CDCRPPDFMSNENTSCIAEIEFLANTPRGIALTFYAVGLTAFVGLVFKRNTP 641
 QY 595 IYRAGGPMCTMLTLLVAVMVVYVGPVKSTCLCRALPCLFTICISIAVBSFQ 654
 Db 642 IYKATNELSYLLFLSLCFSSSLFIEGPDQWTRCLROPAGISFVLCISCLIVYTNR 701
 QY 655 IYCAF--KMASRPRASVWRYGQPVYSAFIVLKWIVYIGMLARPOSHPTDDDP 712
 Db 702 VLVFEAKIPTSFHRK--WMLGLOFLVFLCTFMOILCTIWLTAAPSSYNHHELD 758
 QY 713 KITVSCNPVYRNSLFTNTSL--DLLSVYGFSAFANGKELPYNVEAKFITLSMTFYF 769
 Db 759 EIIPTTC--HEGSLMALGSLIGYTCLLAICFPFAKRSRLPENFNAKFTFSMLIF 815
 QY 770 TTSVSLCTFMSAYSGVLTVYIDLTVNLALISLG---YFGKCYMILFYPERNT 822
 Db 816 IYWIS--FLPVAYSTYGRFVS-AVEYIALAASFGLLACIFENKYYIILFKPSRNT 868

RESULT 15
US-08-480-751-8
Sequence 8, Application US/08480751
Patent No. 5858684
GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,751
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1079 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-751-8

Query Match 24.5%; Score 1088.5; DB 2; Length 1079;
Best Local Similarity 31.7%; Pred. No. 2,2e-94;
Matches 284; Conservative 164; Mismatches 328; Indels 121; Gaps 31.

QY 10 SLFLLWVLAEPNSDFLLP-----GDYLLGLLSIHANMKGIIVHLNLFQVPMCKEY 62

Db 9 ALNALAW-----HSSAYGPDQRAQKKGDITLGGFLPIHF---GVAAKD--QDLAKSRE 56
QY 63 EVKVIQVNL-----MQANFAVEEINNDSSLPGVLLAGEYIVDVC-YISNNVQPVLYELA 116
Db 57 SVECIYRNFRCGRWLOAMIFALEEINSSPLPNNTLGRIPDCNFTVSKALEATLSFVA 116
QY 117 -HEMDLPIQE--DYSNISRNVANIGPDNSQSVTVANFLSLFLPQITYSATISDELRD 173
Db 117 ONKIDSLNDEFCNCESEHPISTIAVAGATGSGVSTAVANLGLFTIPQVSTASSRLLSN 176
QY 174 KYRFPALRTTPSADHIEAMVQMLAFHFMWNIIVLVESDYGDRNGQLGERVARRDIC 233
Db 177 KNQYSFLRTIPNDEHQATAMADIEFRMNMVGIADDDYGRGRIKPREARENERDIC 236
QY 234 IAFQETLPLQPNQMTSEEROLVTYDKLQOASTARVVYVSPDLTIYHFFNEVLRONF 293
Db 237 IDFSLEI-----SOYSDEEIQOVVEVI--ONSTAKYIVVFSSGPDLEPLIKEIVRNRI 288
QY 294 TGAVVIASESNALIDPVNLNLFELGH-LGPFILGTTIOSVPIDPFSEF----- 338
Db 289 TGRIVLASEANASSLI-AMPEYFHVVGCTIGFGAKAQIGFRFELQVHPRKSVHNGF 347
QY 339 -REMGPOA-----GPPPLSRTSOSYTCNQCDCNLTATLSFNTILRLSGE----- 382
Db 348 AKEFWEETFNCHLOAGAGPLPVDFVRSH--EEGNNLNLSSTAFRL--CTGDENINS 403
QY 383 -----RVYVSYSANYAVAAHLSLGC-----DKSTCKRRVYVPMQLEE 423
Db 404 VETPYMDYEHRLISYNYLVAYSINAHLODIYTCPLGRGLFNGSCADIKKEVAQVYK 463
QY 424 IMKVNET-LDHOJFEDDQGVALHLEIVOMQMDNSQNP--FQSVASYPL-----OHOL 475
Db 464 LHLNLTNNMGEOVTFDECGDLYGNTSTINHLSPEDSIVFKEY-GYINNYAKGELF 522
QY 476 KNIDISWHTVNNITPMSCSKRCSGOKRPV-GIHYCECEIDCLPGLNHTEDYE 534
Db 523 INEEKILMSGFREVPFNSCRDQAGTRKGIIEGPCCFECVCPDGEYSGET-DASA 581
QY 535 COACNNNEMSYOSTSCRRKQVLFLMEHAPRIAVALLAAGFLSTALIVYFMHFOTP 594
Db 582 CDKCPDDDFMSENHTSCIAKEIEFLAWTEPRIALTLFAVIGIFLTAFLVGLGFIKFRTP 641
QY 595 IVRSAGPMCFMLTLFLVAVVVPVYGPVSTCLROALFPLCFITICISCIIVRSQ 654
Db 642 IVKATNRELSTLLSLCCRSSSLFFIGEPDMTCRLROPAGISFVLCISCIIVKTR 701
QY 655 IVCAF--KMASRFPRAYSYVWRYGQVYSMAETIVLKVIVYIGMARPOSHRTDPDP 712
Db 702 VLVVEAKIPTSFRHK---WMLNLOFLVFLCTFMQILICIIMLYTAPPSYRNHELED 758
QY 713 KITIVSCPNRNLSLFTSL---DLLSVGFSAVYNGKELPTVYNNAKFTTLEMTYF 769
Db 759 EILFITC---HEGSLMALGSLGYCLLAICFFPAFRKRPENFNNAKFTTSMLETF 815
QY 770 TSSVSLCTFMSAYSGVLTVYDLVTVNLALISG---YEGPCYMLFPEPENT 822
Db 816 IWMIS---FIRAYASTYGRFVS-AVEVAILAASGLACIFPNKYVITILFRPSNFT 868

Search completed: May 19, 2003, 09:53:11
Job time : 18.3799 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:51:25 ; Search time 38.6769 Seconds

(without alignments)
2089.767 Million cell updates/sec

Title: US-09-927-315-9

Perfect score: 4443

Sequence: 1 MGPRKATCSLFFLLWVLAEE.....ERNTPAYENMIGYTRRD 838

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 362588 seqs, 96450795 residues

Total number of hits satisfying chosen parameters: 362588

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_AA:*
1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4443	100.0	838	9	US-09-927-315-9
2	4392.5	98.9	839	9	US-09-897-427A-4
3	4392.5	98.9	839	9	US-10-035-045-21
4	3514	79.1	669	9	US-10-096-144-7
5	3514	79.1	669	12	US-10-124-598-7
6	3203.5	72.1	843	9	US-10-096-144-1
7	3203.5	72.1	843	9	US-09-927-315-7
8	3203.5	72.1	843	12	US-10-124-598-1
9	3151.5	70.9	843	9	US-10-096-144-2
10	3151.5	70.9	843	9	US-09-927-315-8
11	3151.5	70.9	843	12	US-10-124-598-2
12	1454	32.7	840	9	US-09-361-652-1
13	1454	32.7	840	9	US-09-927-315-1
14	1434	32.3	842	9	US-09-361-652-2
15	1434	32.3	842	9	US-09-927-315-2
16	1354	30.5	841	9	US-09-897-427A-2
17	1354	30.5	841	9	US-09-799-629-17
18	1354	30.5	841	9	US-10-035-045-17
19	1354	30.5	841	10	US-09-819-946-2

20	1270	28.6	777	9	US-09-361-652-3	Sequence 3, Appl1
21	1270	28.6	777	9	US-09-927-315-3	Sequence 3, Appl1
22	1267.5	28.5	1138	9	US-10-261-462-4	Sequence 4, Appl1
23	1260.5	28.4	763	10	US-09-819-946-4	Sequence 4, Appl1
24	1094	24.6	1078	12	US-10-002-854-2	Sequence 2, Appl1
25	1093	24.6	1078	10	US-09-727-205-2	Sequence 2, Appl1
26	1092.5	24.6	1027	9	US-10-125-792-2	Sequence 2, Appl1
27	1092.5	24.6	1027	9	US-10-125-792-2	Sequence 2, Appl1
28	1078	24.3	1078	9	US-10-125-792-28	Sequence 28, Appl1
29	1078	24.3	1078	9	US-10-125-792-28	Sequence 28, Appl1
30	1056.5	23.8	941	9	US-10-125-792-8	Sequence 8, Appl1
31	1056.5	23.8	941	9	US-10-125-792-8	Sequence 8, Appl1
32	1049.5	23.6	941	9	US-10-125-792-10	Sequence 10, Appl1
33	1049.5	23.6	941	9	US-10-125-792-10	Sequence 10, Appl1
34	1039	23.4	858	9	US-09-799-629-14	Sequence 14, Appl1
35	1039	23.4	858	9	US-09-927-315-25	Sequence 25, Appl1
36	1039	23.4	858	9	US-10-035-045-14	Sequence 14, Appl1
37	1025.5	23.1	850	9	US-10-125-792-12	Sequence 12, Appl1
38	1025.5	23.1	850	9	US-10-125-792-12	Sequence 12, Appl1
39	1014	22.8	852	9	US-10-261-482-2	Sequence 2, Appl1
40	1014	22.8	852	9	US-10-282-837-14	Sequence 14, Appl1
41	1014	22.8	852	10	US-09-796-338A-14	Sequence 14, Appl1
42	1014	22.8	858	9	US-09-927-315-18	Sequence 18, Appl1
43	1013	22.8	858	9	US-09-927-315-20	Sequence 20, Appl1
44	1012	22.8	858	9	US-09-927-315-23	Sequence 23, Appl1
45	1011	22.8	852	9	US-09-897-427A-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-09-927-315-9

Sequence 9, Application US/09927315
Publication No. US20030040045A1

GENERAL INFORMATION:

APPLICANT: Zuker, Charles S.

APPLICANT: Ryba, Nicholas J.P.

APPLICANT: Nelson, Greg

APPLICANT: Hoon, Mark A.

APPLICANT: Chandrasekar, Jayaram

APPLICANT: Zhang, Yifeng

APPLICANT: The Regents of the University of California

APPLICANT: The Government of the United States of America

APPLICANT: as represented by the Secretary of the

Department of Health and Human Services

TITLE OF INVENTION: Mammalian Sweet Taste Receptors

FILE REFERENCE: 02307E-120110US

CURRENT APPLICATION NUMBER: US/09/927,315

PRIOR FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: US 60/302,898

PRIOR FILING DATE: 2001-07-03

NUMBER OF SEQ ID NOS: 25

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 9

LENGTH: 838

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE: OTHER INFORMATION: human T1R2 sweet taste receptor

US-09-927-315-9

Query Match 100.0%; Score 4443; DB 9; Length 838;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 838; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPRKATCSLFFLLWVLAEEPAENSDFYLPDYLGGFLSHANNKGIYHNLQVPAACK 60
DB 1 MGPRKATCSLFFLLWVLAEEPAENSDFYLPDYLGGFLSHANNKGIYHNLQVPAACK 60
QY 61 EYEKVGIVNLMQAMRAFAVEEINNDSSLLPGVLLGLEYIVDVCYISNNVQPVLYFLAHEDN 120
DB 61 EYEKVGIVNLMQAMRAFAVEEINNDSSLLPGVLLGLEYIVDVCYISNNVQPVLYFLAHEDN 120

```

QY 121 LPIQEDYSNYSIRVAVIGPDNSESVMYANFLSLFLPQITYSASIDELDRKVRPAL 180
DB 121 LPIQEDYSNYSIRVAVIGPDNSESVMYANFLSLFLPQITYSASIDELDRKVRPAL 180
QY 181 LRTTSDADHVEAVOYMLHFRNMWITVYSSDTYGRDNGOLGERRARDICIAFOETL 240
DB 181 LRTTSDADHVEAVOYMLHFRNMWITVYSSDTYGRDNGOLGERRARDICIAFOETL 240
QY 241 PTLQPNQNMNTSEERORLVTYDKLQOSTARVYVFPDLTYHFFNEVLKONFTGAWMIA 300
DB 241 PTLQPNQNMNTSEERORLVTYDKLQOSTARVYVFPDLTYHFFNEVLKONFTGAWMIA 300
QY 301 SESWAIDPVHLNLTGLHGTFLGITTOSVPIPGFSEFERMGPOAGPPPLSRTSOSTCN 360
DB 301 SESWAIDPVHLNLTGLHGTFLGITTOSVPIPGFSEFERMGPOAGPPPLSRTSOSTCN 360
QY 361 QECNCLNATLSEFTILRLSGERVYVSAYAVAAHLSLGCDSCTCKRKYVPMOL 420
DB 361 QECNCLNATLSEFTILRLSGERVYVSAYAVAAHLSLGCDSCTCKRKYVPMOL 420
QY 421 LEEIMKVNFTLLDHOIFEDPOGDVALHLEIYQOMQDRSONPQOSVASYPLQROLKNIQD 480
DB 421 LEEIMKVNFTLLDHOIFEDPOGDVALHLEIYQOMQDRSONPQOSVASYPLQROLKNIQD 480
QY 481 ISWHTVNTTIPMSCKRQSGOKKRPVGIHCCFECIDCLPGTFNLHTEDEYECQACP 540
DB 481 ISWHTVNTTIPMSCKRQSGOKKRPVGIHCCFECIDCLPGTFNLHTEDEYECQACP 540
QY 541 NEMYSOSETSCFRQOLVFLHMEHAPTAVALALAGLSTALIVYWRHQTIVYSAG 600
DB 541 NEMYSOSETSCFRQOLVFLHMEHAPTAVALALAGLSTALIVYWRHQTIVYSAG 600
QY 601 GPMCFMLTLLVAVYVYVGPVKVSTCLCROALPPLCTTICISCIAYRSQIYCAFK 660
DB 601 GPMCFMLTLLVAVYVYVGPVKVSTCLCROALPPLCTTICISCIAYRSQIYCAFK 660
QY 661 MASFPRAYSYWRVYQGPYVSMATVYKAVIYVIGMLARPOSHPRTDDPKITIVSCN 720
DB 661 MASFPRAYSYWRVYQGPYVSMATVYKAVIYVIGMLARPOSHPRTDDPKITIVSCN 720
QY 721 NPNYRNSLLENTSLDLLSVGFSAVMGKELPTVYNEAKFTLSMTYFTSSVSLCTFM 780
DB 721 NPNYRNSLLENTSLDLLSVGFSAVMGKELPTVYNEAKFTLSMTYFTSSVSLCTFM 780
QY 781 AYSGLVTVTDLVTLVNLALISLGYGPKCYMILFPERNTPAYFNSMIOGYTMRD 838
DB 781 AYSGLVTVTDLVTLVNLALISLGYGPKCYMILFPERNTPAYFNSMIOGYTMRD 838

RESULT 2
US-09-897-427A-4
; Sequence 4, Application US/09897427A
; Patent No. US20020160424A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIADONG
; APPLICANT: STASZEWSKI, LENA
; APPLICANT: XU, HONG
; APPLICANT: EHEVERRI, FERNANDO
; TITLE OF INVENTION: T1R HETERO-OLIGOMERIC TASTE RECEPTORS
; FILE REFERENCE: 078003-0282558
; CURRENT APPLICATION NUMBER: US/09/897,427A
; CURRENT FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 839
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-427A-4
Query Match 98.9%; Score 4392.5; DB 9; Length 839;

```

```

Best Local Similarity 99.3%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 MGPRAKTCISLFLWLTAPAEKSDPYLPQDYLLGGLFSLHNMKIVLNLQVPMCK 60
DB 1 MGPRAKTCISLFLWLTAPAEKSDPYLPQDYLLGGLFSLHNMKIVLNLQVPMCK 60
QY 61 EYEKVIYGNLMQAMRAVEEINNDSSLPGVLLGYEIVVCYISNNVQVLYFLAHEDN 120
DB 61 EYEKVIYGNLMQAMRAVEEINNDSSLPGVLLGYEIVVCYISNNVQVLYFLAHEDN 120
QY 121 LPIQEDYSNYSIRVAVIGPDNSESVMYANFLSLFLPQITYSASIDELDRKVRPAL 180
DB 121 LPIQEDYSNYSIRVAVIGPDNSESVMYANFLSLFLPQITYSASIDELDRKVRPAL 180
QY 181 LRTTSDADHVEAVOYMLHFRNMWITVYSSDTYGRDNGOLGERRARDICIAFOETL 240
DB 181 LRTTSDADHVEAVOYMLHFRNMWITVYSSDTYGRDNGOLGERRARDICIAFOETL 240
QY 241 PTLQPNQNMNTSEERORLVTYDKLQOSTARVYVFPDLTYHFFNEVLKONFTGAWMIA 300
DB 241 PTLQPNQNMNTSEERORLVTYDKLQOSTARVYVFPDLTYHFFNEVLKONFTGAWMIA 300
QY 301 SESWAIDPVHLNLTGLHGTFLGITTOSVPIPGFSEFERMGPOAGPPPLSRTSOSTCN 360
DB 301 SESWAIDPVHLNLTGLHGTFLGITTOSVPIPGFSEFERMGPOAGPPPLSRTSOSTCN 360
QY 361 QECNCLNATLSEFTILRLSGERVYVSAYAVAAHLSLGCDSCTCKRKYVPMOL 420
DB 361 QECNCLNATLSEFTILRLSGERVYVSAYAVAAHLSLGCDSCTCKRKYVPMOL 420
QY 421 LEEIMKVNFTLLDHOIFEDPOGDVALHLEIYQOMQDRSONPQOSVASYPLQROLKNIQD 480
DB 421 LEEIMKVNFTLLDHOIFEDPOGDVALHLEIYQOMQDRSONPQOSVASYPLQROLKNIQD 480
QY 481 ISWHTVNTTIPMSCKRQSGOKKRPVGIHCCFECIDCLPGTFNLHTEDEYECQACP 540
DB 481 ISWHTVNTTIPMSCKRQSGOKKRPVGIHCCFECIDCLPGTFNLHTEDEYECQACP 540
QY 541 NEMYSOSETSCFRQOLVFLHMEHAPTAVALALAGLSTALIVYWRHQTIVYSAG 600
DB 541 NEMYSOSETSCFRQOLVFLHMEHAPTAVALALAGLSTALIVYWRHQTIVYSAG 600
QY 601 GPMCFMLTLLVAVYVYVGPVKVSTCLCROALPPLCTTICISCIAYRSQIYCAFK 660
DB 601 GPMCFMLTLLVAVYVYVGPVKVSTCLCROALPPLCTTICISCIAYRSQIYCAFK 660
QY 661 MASFPRAYSYWRVYQGPYVSMATVYKAVIYVIGMLARPOSHPRTDDPKITIVSC 719
DB 661 MASFPRAYSYWRVYQGPYVSMATVYKAVIYVIGMLARPOSHPRTDDPKITIVSC 719
QY 720 NPNYRNSLLENTSLDLLSVGFSAVMGKELPTVYNEAKFTLSMTYFTSSVSLCTFM 779
DB 720 NPNYRNSLLENTSLDLLSVGFSAVMGKELPTVYNEAKFTLSMTYFTSSVSLCTFM 779
QY 781 AYSGLVTVTDLVTLVNLALISLGYGPKCYMILFPERNTPAYFNSMIOGYTMRD 838
DB 781 AYSGLVTVTDLVTLVNLALISLGYGPKCYMILFPERNTPAYFNSMIOGYTMRD 838

RESULT 3
US-10-035-045-21
; Sequence 21, Application US/10035045
; Publication No. US20030054448A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIADONG
; APPLICANT: STASZEWSKI, LENA
; APPLICANT: O'CONNELL, SHAWN
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: T1R TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003-0280681
; CURRENT APPLICATION NUMBER: US/10/035, 045

```

; CURRENT FILING DATE: 2002-01-03
 ; PRIOR APPLICATION NUMBER: 60/259,227
 ; PRIOR FILING DATE: 2001-01-03
 ; PRIOR APPLICATION NUMBER: 60/284,547
 ; PRIOR FILING DATE: 2001-04-19
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 21
 ; LENGTH: 839
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-035-045-21

Query Match 98.9%; Score 4392.5; DB 9; Length 839;
 Best Local Similarity 99.3%; Pred. No. 0;

Matches 833; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 MGPRAKTSLEFLLVLAEPANSDPYLPDYLGLGSLHANMGIHLNQLQVPMCK 60
 DB 1 MGPRAKTSLEFLLVLAEPANSDPYLPDYLGLGSLHANMGIHLNQLQVPMCK 60
 QY 61 EYEVKIVGYNLQAMRFAVEEINNDSSLPGVLLGYEIVDVCYISNNVQPVLYFLAHEDN 120
 DB 61 EYEVKIVGYNLQAMRFAVEEINNDSSLPGVLLGYEIVDVCYISNNVQPVLYFLAHEDN 120
 QY 121 LPIQEDYNTISRVAIVGIPDNSESMTVANFLSLFLPQITYSASIDELRDKVAREPAL 180
 DB 121 LPIQEDYNTISRVAIVGIPDNSESMTVANFLSLFLPQITYSASIDELRDKVAREPAL 180
 QY 181 LRTTSPADHHEAVQMLHFRMNIIVLVSSDPTVGRDNGOLLGERVARDCIAFOETL 240
 DB 181 LRTTSPADHHEAVQMLHFRMNIIVLVSSDPTVGRDNGOLLGERVARDCIAFOETL 240
 QY 241 PTLQPNQNTSEERQRLVTVKLOQSTARVYVESPDLTLHFENEVLRONFTGAWMTA 300
 DB 241 PTLQPNQNTSEERQRLVTVKLOQSTARVYVESPDLTLHFENEVLRONFTGAWMTA 300
 QY 301 SESMAIDPVLAHLTELGHGTFELITIQSVPIPGFSEFREMGPQAPPLSRTSOSTCN 360
 DB 301 SESMAIDPVLAHLTELGHGTFELITIQSVPIPGFSEFREMGPQAPPLSRTSOSTCN 360
 QY 361 QECDCNLANLTSFNTILRLSGERVVSVYSAVVAHALHSLGCDKSTCTKRVVPMOL 420
 DB 361 QECDCNLANLTSFNTILRLSGERVVSVYSAVVAHALHSLGCDKSTCTKRVVPMOL 420
 QY 421 LEEIMKVNFTLDHOFEPDQGVALLHETIYOMQDRSONPQOSVASYPLQRLKNID 480
 DB 421 LEEIMKVNFTLDHOFEPDQGVALLHETIYOMQDRSONPQOSVASYPLQRLKNID 480
 QY 481 ISMHTVNTIIPMSCKRQSGOKKRPVGIHVCCFECIDCLPCTFLNHTEDEECQACPN 540
 DB 481 ISMHTVNTIIPMSCKRQSGOKKRPVGIHVCCFECIDCLPCTFLNHTEDEECQACPN 540
 QY 541 NEMTSQSESCFRKQLVLEHHEAPITAVALLAAGLSTLAILVTFMRHPTPIVRSAG 600
 DB 541 NEMTSQSESCFRKQLVLEHHEAPITAVALLAAGLSTLAILVTFMRHPTPIVRSAG 600
 QY 601 GPMCFMLTLLVAVYVYVGPVSTCLCROALFPLCTITCSCIIVRSQIYCAK 660
 DB 601 GPMCFMLTLLVAVYVYVGPVSTCLCROALFPLCTITCSCIIVRSQIYCAK 660
 QY 661 MASRPRAVSVMRYOGPVSMATIVLKMVIVIGMLARPOS-HPRTPDDPKITIVSC 719
 DB 661 MASRPRAVSVMRYOGPVSMATIVLKMVIVIGMLARPOS-HPRTPDDPKITIVSC 719
 QY 720 NENYRNSLTFNTSLDLSVGFSPAYMGKELPTNYNEAKFTTSLMTFTSVSLCTEM 779
 DB 720 NENYRNSLTFNTSLDLSVGFSPAYMGKELPTNYNEAKFTTSLMTFTSVSLCTEM 779
 QY 780 SAYSGVLTIVDLTVLNLALISLGYFPGPKCYMLIFPERMTPAVENMIGGYMRD 838
 DB 780 SAYSGVLTIVDLTVLNLALISLGYFPGPKCYMLIFPERMTPAVENMIGGYMRD 838

RESULT 4
 ; US-10-096-144-7
 ; Sequence 7; Application US/10096144
 ; Publication No. US20030022288A1
 ; GENERAL INFORMATION:

; APPLICANT: Zuker, Charles S.
 ; APPLICANT: Adler, Jon Elliott
 ; APPLICANT: Lindemeier, Juergen
 ; APPLICANT: The Regents of the University of California
 ; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
 ; FILE REFERENCE: 02307E-088720US
 ; CURRENT APPLICATION NUMBER: US/10/096,144

; CURRENT FILING DATE: 2002-03-08
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/361,631
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/095,464
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747

; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 7

; LENGTH: 669
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens

; OTHER INFORMATION: human G-protein coupled receptor (GPCR) B4 amino
 ; OTHER INFORMATION: acid sequence
 ; US-10-096-144-7

Query Match 79.1%; Score 3514; DB 9; Length 669;
 Best Local Similarity 98.4%; Pred. No. 3,6e-288;

Matches 666; Conservative 1; Mismatches 2; Indels 8; Gaps 2;

QY 162 ITYSASIDELRDKVAREPALRTTSPADHHEAVQMLHFRMNIIVLVSSDPTVGRDNG 221
 DB 1 ITYSASIDELRDKVAREPALRTTSPADHHEAVQMLHFRMNIIVLVSSDPTVGRDNG 221
 QY 222 LIGERVARRDICIATFOETLPTLQPNQNTSEERQRLVTVKLOQSTARVYVESPDLTL 281
 DB 222 LIGERVARRDICIATFOETLPTLQPNQNTSEERQRLVTVKLOQSTARVYVESPDLTL 281
 QY 282 YHFEENEVLRONFTGAWMTASEMAIDPVLAHLTELGHGTFELITIQSVPIPGFSEFREM 341
 DB 282 YHFEENEVLRONFTGAWMTASEMAIDPVLAHLTELGHGTFELITIQSVPIPGFSEFREM 341
 QY 342 GPQAGPPPLSRTSOSTCNQECDCNLANLTSFNTILRLSGERVVSVYSAVVAHALHS 401
 DB 342 GPQAGPPPLSRTSOSTCNQECDCNLANLTSFNTILRLSGERVVSVYSAVVAHALHS 401
 QY 402 LIGCDKSTCTKRVVPMOLLEIIMKVNFTLDHOFEPDQGVALLHETIYOMQDRSONP 461
 DB 402 LIGCDKSTCTKRVVPMOLLEIIMKVNFTLDHOFEPDQGVALLHETIYOMQDRSONP 461
 QY 462 FOSVASYPLQRLKNIDISMTVNTIIPMSCKRQSGOKKRPVGIHVCCFECIDCL 521
 DB 462 FOSVASYPLQRLKNIDISMTVNTIIPMSCKRQSGOKKRPVGIHVCCFECIDCL 521
 QY 522 PGTFLNHTEDEECQACPNNEMTSQSESCFRKQLVLEHHEAPITAVALLAAGLSTL 581
 DB 522 PGTFLNHTEDEECQACPNNEMTSQSESCFRKQLVLEHHEAPITAVALLAAGLSTL 581
 QY 582 ALIVTFMRHPTPIVRSAGGPMCFMLTLLVAVYVYVGPVSTCLCROALPLCF 641
 DB 582 ALIVTFMRHPTPIVRSAGGPMCFMLTLLVAVYVYVGPVSTCLCROALPLCF 641
 QY 642 TTCISCIIVRSQIYCAFMASRPRAVSVMRYOGPVSMATIVLKMVIVIGMLARP 701
 DB 642 TTCISCIIVRSQIYCAFMASRPRAVSVMRYOGPVSMATIVLKMVIVIGMLARP 701
 QY 702 QSHPRTPDDPKITIVSCPNYRNSLTFNTSLDLSVGFSPAYMGKELPTNYNEAKFT 761

|||||
Db 533 QSHRPDPDPKTIIVSCNPNYRNSLFFNTSLDLLSVGFSPAYMKELPTYNNEAKFI 592
QY 762 TLSMTFFYTSVSLCTFMSAYSGVLVTYIDLTVNLALISLGYFGPKCYMILFPERN 821
Db 593 TLSMTFFYTSVSLCTFMSAYSGVLVTYIDLTVNLALISLGYFGPKCYMILFPERN 652
QY 822 TPAYFNSMIGYTMRRD 838
Db 653 TPAYFNSMIGYTMRRD 669

RESULT 5
US-10-124-598-7
Sequence 7, Application US/10124598
Patent No. US20020119526A1

GENERAL INFORMATION:
APPLICANT: Zuker, Charles S.
APPLICANT: Adler, Jon Elliot
APPLICANT: Lindemeier, Juergen
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: The Regents of the University of California
TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
TITLE OF INVENTION: Involved in Sensory Transduction
FILE REFERENCE: 02307E-088720US
CURRENT APPLICATION NUMBER: US/10/124, 598
CURRENT FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/361, 631
PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112, 747
PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 669
TYPE: PRF
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human G-protein coupled receptor (GPCR) B4 amino
US-10-124-598-7

Query Match 79.1%; Score 3514; DB 12; Length 669;
Best Local Similarity 98.4%; Pred. No. 3.6e-288;
Matches 666; Conservative 1; Mismatches 2; Indels 8; Gaps 2;

QY 162 ITYSALSDERLDRKVRFPALRTTPSADHVEAMVQMLHFRMNIIVLVSSDYGRNGQ 221
Db 1 ITYSALSDERLDRKVRFPALRTTPSADHVEAMVQMLHFRMNIIVLVSSDYGRNGQ 60
QY 222 LIGERRARDICIAFOETPLTOPNOMTSEEROLVTYIDKLOQSTARVAVVSPDLTL 281
Db 61 LIGERRARDICIAFOETPLTOPNOMTSEEROLVTYIDKLOQSTARVAVVSPDLTL 120
QY 282 YHFEENEVLRONFTGAVVIAESMAIDVYLNHLELGHGFLGTLTISVPIPGSEEREW 341
Db 121 YHFEENEVLRONFTGAVVIAESMAIDVYLNHLELGHGFLGTLTISVPIPGSEEREW 180
QY 342 GPQAGPPPLSRSTQSYTCNOCNCLNATLSFTITLRLSGERVVSYSAVVAHAHLS 401
Db 181 GPQAGPPPLSRSTQSYTCNOCNCLNATLSFTITLRLSGERVVSYSAVVAHAHLS 240
QY 402 LIGCDKSTCKRVVYPMQLEELIKVNFLLDHOIFPDPOGDVALHLEIYQWMDRSQNP 461
Db 241 LIGCDKSTCKRVVYPMQLEELIKVNFLLDHOIFPDPOGDVALHLEIYQWMDRSQNP 300
QY 462 FQSYASTYPIQROLKNTQDISMTVNTTIPMSCKSKRCQSGQKKRPVGIHVCCFECIDCL 521
Db 301 FQSYASTYPIQROLKNTQDISMTVNTTIPMSCKSKRCQSGQKKRPVGIHVCCFECIDCL 359
QY 522 PGFTLNTEDYEQACPNMNSYQSETSCFKROLVFLWHEAPTIVALLAALGFTSL 581
Db 360 PGFTLNTEDYEQACPNMNSYQSETSCFKROLVFLWHEAPTIVALLAALGFTSL 412

QY 582 ALIVFMRHQPPIVRSAGSPKCFMLTLLVAVVWVYVGPVKSTCICRQALPELPCF 641
Db 413 ALIVFMRHQPPIVRSAGSPKCFMLTLLVAVVWVYVGPVKSTCICRQALPELPCF 472
QY 642 TICISCIASVSPQIVCAFKNASRPFRAYSYWYRQGYVSMATFYUKMYIYIGMARP 701
Db 473 TICISCIASVSPQIVCAFKNASRPFRAYSYWYRQGYVSMATFYUKMYIYIGMARP 532
QY 702 QSHRPDPDPKTIIVSCNPNYRNSLFFNTSLDLLSVGFSPAYMKELPTYNNEAKFI 761
Db 533 QSHRPDPDPKTIIVSCNPNYRNSLFFNTSLDLLSVGFSPAYMKELPTYNNEAKFI 592
QY 762 TLSMTFFYTSVSLCTFMSAYSGVLVTYIDLTVNLALISLGYFGPKCYMILFPERN 821
Db 593 TLSMTFFYTSVSLCTFMSAYSGVLVTYIDLTVNLALISLGYFGPKCYMILFPERN 652
QY 822 TPAYFNSMIGYTMRRD 838
Db 653 TPAYFNSMIGYTMRRD 669

RESULT 6
US-10-096-144-1
Sequence 1, Application US/10096144
Publication No. US2003002288A1

GENERAL INFORMATION:
APPLICANT: Zuker, Charles S.
APPLICANT: Adler, Jon Elliot
APPLICANT: Lindemeier, Juergen
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
TITLE OF INVENTION: Involved in Sensory Transduction
FILE REFERENCE: 02307E-088720US
CURRENT APPLICATION NUMBER: US/10/096, 144
CURRENT FILING DATE: 2002-03-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/361, 631
PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/095, 464
PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112, 747
PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 843
TYPE: PRF
ORGANISM: Rattus sp.
FEATURE:
OTHER INFORMATION: rat G-protein coupled receptor (GPCR) B4 amino
US-10-096-144-1

Query Match 72.1%; Score 3203.5; DB 9; Length 843;
Best Local Similarity 70.6%; Pred. No. 8.3e-262;
Matches 595; Conservative 108; Mismatches 133; Indels 7; Gaps 4;

QY 1 MGPRAKTCISLFLWLVAEP--AENSDFYLPQDYLGLSLHANKGIVHLNFIQVP 57
Db 1 MGPRAKTCISLFLWLVAEP--AENSDFYLPQDYLGLSLHANKGIVHLNFIQVP 60
QY 58 MCKEYEVKIVGYNLMQAMRAVEEINNDSSILPVLGLEYIVDYCYISNNVQPLTYFLAH 117
Db 61 MCKEYEVKIVGYNLMQAMRAVEEINNDSSILPVLGLEYIVDYCYISNNVQPLTYFLAH 120
QY 118 EDNLPLIQEDYSYNTISVVAVVIGPDNSYWTYANFLSLFLPQITYSALSDERLDRKVR 177
Db 121 DDLLPLIKDYSOYMPHVAVVIGPDNSYWTYANFLSLFLPQITYSALSDERLDRKVR 180
QY 178 PALIRTPPSADHVEAMVQMLHFRMNIIVLVSSDYGRNGQALGGERAR-RDICIACF 236
Db 181 PSMKRTYPSATHTHEANVQMLHFRMNIIVLVSSDYGRNGQALGGERAR-RDICIACF 240
QY 237 QETPLPLQPNOMNTSEEROLVTYIDKLOQSTARVAVVSPDLTLVYHFEENEVLRONFTGA 296

```
Db 241 QEVLLIPSSQVMSRBEQROLDNIIDKLRTSARVYVSPSLXSFHEVLKRNFTG 300
Oy 237 VWIASENAIDVYLNHLTELGLTFLGTTIOSVPIPGSEFEREMGPQAGPPLSRTSOS 356
Db 301 VWIASENAIDVYLNHLTELGLTFLGTTIOSVPIPGSEFEREMGPQAGPPLSRTSOS 360
Oy 357 YTCNOECNCLNATLSEFTILRLSGERYVYSAYVAVAHLSLGLGDKSTCRKRYV 416
Db 361 TTCNODCACLNTTTSFNNILLSGERVYYSAYVAVAHLSLGLGDKSTCRKRYV 420
Oy 417 PMQLLEIKVNFETLDHIFEDPOGDAVLAHEIYQWMDRSONPQOVSASTYPILORLK 476
Db 421 PMQLLEIKVNFETLDHIFEDPOGDAVLAHEIYQWMDRSONPQOVSASTYPILORLK 480
Oy 477 NIQDISMHTVNTTIPMSKSRKQSGOKKRPVGHVCEGIDCLPGFTLNTDEBEYEQ 536
Db 481 YINNVSWITPNTVPMVSMKSCQOPQMKSVGLHPCCECLDMPGTYLNSADEFNCL 540
Oy 537 ACPNNEMSYOSETSCFKROLVLEWHEAPTIVALLAAGFLSTLAILYFMRHPTPIV 596
Db 541 SCPGSMKSKNDITCFQRPTFLHMEVPTIVALLAAGFLSTLAILYFMRHPTPIV 600
Oy 597 RSAGPMPCLMTLLVAVMYVYVGPVKYSTCLRQALFPLCTICISCIANSFOY 656
Db 601 RSAGPMPCLMTLLVAVMYVYVGPVKYSTCLRQALFPLCTICISCIANSFOY 660
Oy 657 CAFKASRPRAYSVWVRQGPVYSMAFTVUKMYIVVIGMLARQSHB--RTDDDPKI 714
Db 661 CVFKARRLPSAYSFMMRHGYPVFAFTAKVALVGNMMA--TTINIGRTDDDPKI 719
Oy 715 YTCNOECNCLNATLSEFTILRLSGERYVYSAYVAVAHLSLGLGDKSTCRKRYV 774
Db 720 MTLSCHPNTRNCLNTSMDLLSVLGSFAFMGKELPNYNEAKFTITSMFTSSIS 779
Oy 775 LCTFMSAIVGLVTVLTVLNLIAISLGFSPKCYMILFYPERNTPAYNSMIQGYT 834
Db 780 LCTFMSVHDGVLTIMDLVTVLNLIAISLGFSPKCYMILFYPERNTPAYNSMIQGYT 839
Oy 835 MRR 837
Db 840 MRR 842

RESULT 7
US-09-927-315-7
; Sequence 7, Application US/09927315
; Publication No. US20030040045A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nicholas J.P.
; APPLICANT: Nelson, Greg
; APPLICANT: Hoon, Mark A.
; APPLICANT: Chandrasekar, Jayaram
; APPLICANT: Zhang, Yifeng
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
; FILE REFERENCE: 02307E-120110US
; CURRENT APPLICATION NUMBER: US/09/927,315
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/302,898
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: rat Tir2 sweet taste receptor
```

```
US-09-927-315-7
Query Match 72.1%; Score 3203.5; DB 9; Length 843;
Best Local Similarity 70.6%; Pred. No. 8.3e-262;
Matches 595; Conservative 108; Mismatches 133; Indels 7; Gaps 4;

Oy 1 MGPRAKTICSLFELLAVLAEP---AENSDFYLPQDVLGGLSFLANMKMGYHLNLOVP 57
Db 1 MGPRAKTICSLFELLAVLAEP---AENSDFYLPQDVLGGLSFLANMKMGYHLNLOVP 60
Oy 58 MCKEYEVKIGYVNLQAMFAVEEINNDSSLLPGVLLGEIYDVCYISNNVQVYFLAH 117
Db 61 KCEFTMKVLYGNLQAMFAVEEINNDSSLLPGVLLGEIYDVCYISNNVQVYFLAH 120
Oy 118 EDNLPIQEDYNSYRYAVYIGPDSSEVMYANFLSFLPOTYSAISDELDRKYE 177
Db 121 DDLLPIKLDYQYMHVAVYIGPDSSEVMYANFLSFLPOTYSAISDELDRKYE 180
Oy 178 PALRTTSPADHVEAMVOLMIFRNMWIIYVSSDPTGROGQLGEERVAR-RDICIAP 236
Db 181 PSLMTATVSATIHIEAMVOLMIFRNMWIIYVSSDPTGROGQLGEERVAR-RDICIAP 240
Oy 237 QETLPTLPQNMNTSEEROLYTVDKLQOSTARVYVSPDLTYHFENEVLKRNFTGA 296
Db 241 QEVLLIPSSQVMSRBEQROLDNIIDKLRTSARVYVSPSLXSFHEVLKRNFTG 300
Oy 297 VWIASENAIDVYLNHLTELGLTFLGTTIOSVPIPGSEFEREMGPQAGPPLSRTSOS 356
Db 301 VWIASENAIDVYLNHLTELGLTFLGTTIOSVPIPGSEFEREMGPQAGPPLSRTSOS 360
Oy 357 YTCNOECNCLNATLSEFTILRLSGERYVYSAYVAVAHLSLGLGDKSTCRKRYV 416
Db 361 TTCNODCACLNTTTSFNNILLSGERVYYSAYVAVAHLSLGLGDKSTCRKRYV 420
Oy 421 PMQLLEIKVNFETLDHIFEDPOGDAVLAHEIYQWMDRSONPQOVSASTYPILORLK 476
Db 477 NIQDISMHTVNTTIPMSKSRKQSGOKKRPVGHVCEGIDCLPGFTLNTDEBEYEQ 536
Oy 477 NIQDISMHTVNTTIPMSKSRKQSGOKKRPVGHVCEGIDCLPGFTLNTDEBEYEQ 536
Db 481 YINNVSWITPNTVPMVSMKSCQOPQMKSVGLHPCCECLDMPGTYLNSADEFNCL 540
Oy 537 ACPNNEMSYOSETSCFKROLVLEWHEAPTIVALLAAGFLSTLAILYFMRHPTPIV 596
Db 541 SCPGSMKSKNDITCFQRPTFLHMEVPTIVALLAAGFLSTLAILYFMRHPTPIV 600
Oy 597 RSAGPMPCLMTLLVAVMYVYVGPVKYSTCLRQALFPLCTICISCIANSFOY 656
Db 601 RSAGPMPCLMTLLVAVMYVYVGPVKYSTCLRQALFPLCTICISCIANSFOY 660
Oy 657 CAFKASRPRAYSVWVRQGPVYSMAFTVUKMYIVVIGMLARQSHB--RTDDDPKI 714
Db 661 CVFKARRLPSAYSFMMRHGYPVFAFTAKVALVGNMMA--TTINIGRTDDDPKI 719
Oy 715 YTCNOECNCLNATLSEFTILRLSGERYVYSAYVAVAHLSLGLGDKSTCRKRYV 774
Db 720 MTLSCHPNTRNCLNTSMDLLSVLGSFAFMGKELPNYNEAKFTITSMFTSSIS 779
Oy 775 LCTFMSAIVGLVTVLTVLNLIAISLGFSPKCYMILFYPERNTPAYNSMIQGYT 834
Db 780 LCTFMSVHDGVLTIMDLVTVLNLIAISLGFSPKCYMILFYPERNTPAYNSMIQGYT 839
Oy 835 MRR 837
Db 840 MRR 842

RESULT 8
US-10-124-598-1
; Sequence 1, Application US/10124598
; Patent No. US20020119526A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
```

```

; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeyer, Juergen
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-protein Coupled Receptor
; TITLE OF INVENTION: Involved in Sensory Transduction
; FILE REFERENCE: 02307E-08872005
; CURRENT APPLICATION NUMBER: US/10/124,598
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/361,631
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 843
; TYPE: PR
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: rat G-protein coupled receptor (GPCR) B4 amino
; OTHER INFORMATION: acid sequence
US-10-124-598-1

```

```

Query Match      72.1%; Score 3203.5; DB 12; Length 843;
Best Local Similarity 70.6%; Pred. No. 8.3e-262;
Matches 595; Conservative 108; Mismatches 133; Indels 7; Gaps 4;

```

```

QY 1 MGPAKATCSLFFLLVLAEP---AENSDFYLPDGYLLGGLFSLHANKGIYHLNFIQVP 57
D 1 MGPAKATCSLFFLLVLAEP---AENSDFYLPDGYLLGGLFSLHANKGIYHLNFIQVP 60
QY 58 MCKEYKAVIGYINLMQAMRAVEEINNDSSLLPGLVLYGYEIVDVCYISNNVQVLYFLAH 117
D 61 KCEFTKAVIGYINLMQAMRAVEEINNDSSLLPGLVLYGYEIVDVCYISNNVQVLYFLAH 120
QY 118 EDNLPIQEDSNYSISRYAVIIGPDSSEVMTANFLSLFLPQITYSALDELDRDYRF 177
D 121 DDLLPLIKDYSQYMPHVAIVIGPDSSEVMTANFLSLFLPQITYSALDELDRDYRF 180
QY 178 PALRTTPSADHVEAVQVLAHFRWNIIVLVSSDYGRDNGOLGERVAR-RDICTAF 236
D 181 PSMATVPANHHIEAVQVLAHFRWNIIVLVSSDYGRDNGOLGERVAR-RDICTAF 240
QY 237 QETPLTQPNOMNTSEERQRLVTVDKLOOSTARVVVSPDLTLVYFNEVLRQNTGA 296
D 241 QEVLPPESSQVMSERQRLVTVDKLOOSTARVVVSPDLTLVYFNEVLRQNTGA 300
QY 297 VWIASEMAIDPVLAHNLTELHGTFLGVTIQRSVIFGSEFRERMGQAPPLSRQSOS 356
D 301 VWIASEMAIDPVLAHNLTELHGTFLGVTIQRSVIFGSEFRERMGQAPPLSRQSOS 360
QY 357 YTCNOECDNATLSEFNTILRLSGERVVSVYSAVAVAHLSLLGCDSTCTKRVY 416
D 361 TTCNOECDNATLSEFNTILRLSGERVVSVYSAVAVAHLSLLGCDSTCTKRVY 420
QY 417 PMOLLEIKVYNTFLDHOIFPDGQVALHLEIVQWMDRSONPDSVASVYPLQRLK 476
D 421 PMOLLEIKVYNTFLDHOIFPDGQVALHLEIVQWMDRSONPDSVASVYPLQRLK 480
QY 477 NIODISHTVNTIPMSCKRCSGOKKRPVGIHVCCECTIDCLPGFLNHTDEECQ 536
D 481 YINNVSWTPNNTVPMSCSKSCOPGOMKKSQVGLHPCFECCLDMPGTILNRSADERNCL 540
QY 537 APCNNESYOSSETSCFQRLVLEHMEAPTLAVALAAGLSTLALIVFWRHFOPTIV 596
D 541 SCPSMSYKNDICFCORBPFLFEMHEVPTIVAIALGFSFLALIFEMRHFQPMV 600
QY 597 RSAGGPCFLLMLLVAIVMYVYVGPVKYSTLCROALPFLCFTICISIANVRSQIV 656
D 601 RSAGGPCFLLMLLVAIVMYVYVGPVKYSTLCROALPFLCFTICISIANVRSQIV 660
QY 657 CAFEMASRPFAYSYVWRYOGPYVSMAPFVYLVKVIIVVIGLARPOSHP--RTDPDPKI 714
D 657 CAFEMASRPFAYSYVWRYOGPYVSMAPFVYLVKVIIVVIGLARPOSHP--RTDPDPKI 714

```

```

D 661 CVFKMARLPSAVSFMMRKHGPVYVAFITAKVALVGNMIA-TTINPIGRDDDDPNI 719
QY 715 TIVSCNPNRYNSLLENTSLDLSVGFSPAYMKEPLTNYNEAKFTILSMTEFFTSVS 774
D 720 MTLSCNPNRYNSLLENTSLDLSVGFSPAYMKEPLTNYNEAKFTILSMTEFFTSVS 779
QY 775 LCTFMSAYSGVLYTVLTVLNLALISLGFPGKCYMILFEYERNTPAVFNMIQGT 834
D 780 LCTFMSAYSGVLYTVLTVLNLALISLGFPGKCYMILFEYERNTPAVFNMIQGT 839
QY 835 MR 837
D 840 MR 842

```

```

RESULT 9
US-10-096-144-2
; Sequence 2, Application US/10096144
; Publication No. US2003002288A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeyer, Juergen
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-protein Coupled Receptor
; TITLE OF INVENTION: Involved in Sensory Transduction
; FILE REFERENCE: 02307E-08872005
; CURRENT APPLICATION NUMBER: US/10/096,144
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/361,631
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/095,464
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 843
; TYPE: PR
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) B4 amino
; OTHER INFORMATION: acid sequence
US-10-096-144-2

```

```

Query Match      70.9%; Score 3151.5; DB 9; Length 843;
Best Local Similarity 68.9%; Pred. No. 2e-257;
Matches 581; Conservative 113; Mismatches 142; Indels 7; Gaps 4;

```

```

QY 1 MGPAKATCSLFFLLVLAEP---AENSDFYLPDGYLLGGLFSLHANKGIYHLNFIQVP 57
D 1 MGPAKATCSLFFLLVLAEP---AENSDFYLPDGYLLGGLFSLHANKGIYHLNFIQVP 60
QY 58 MCKEYKAVIGYINLMQAMRAVEEINNDSSLLPGLVLYGYEIVDVCYISNNVQVLYFLAH 117
D 61 KCEFTKAVIGYINLMQAMRAVEEINNDSSLLPGLVLYGYEIVDVCYISNNVQVLYFLAH 120
QY 118 EDNLPIQEDSNYSISRYAVIIGPDSSEVMTANFLSLFLPQITYSALDELDRDYRF 177
D 121 DDLLPLIKDYSQYMPHVAIVIGPDSSEVMTANFLSLFLPQITYSALDELDRDYRF 180
QY 178 PALRTTPSADHVEAVQVLAHFRWNIIVLVSSDYGRDNGOLGERVAR-RDICTAF 236
D 181 PSMATVPANHHIEAVQVLAHFRWNIIVLVSSDYGRDNGOLGERVAR-RDICTAF 240
QY 237 QETPLTQPNOMNTSEERQRLVTVDKLOOSTARVVVSPDLTLVYFNEVLRQNTGA 296
D 241 QEVLPPESSQVMSERQRLVTVDKLOOSTARVVVSPDLTLVYFNEVLRQNTGA 300
QY 297 VWIASEMAIDPVLAHNLTELHGTFLGVTIQRSVIFGSEFRERMGQAPPLSRQSOS 356
D 301 VWIASEMAIDPVLAHNLTELHGTFLGVTIQRSVIFGSEFRERMGQAPPLSRQSOS 360

```


; CURRENT APPLICATION NUMBER: US/10/124, 598
 ; CURRENT FILING DATE: 2002-04-16
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/361, 631
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-27
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112, 747
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 843
 ; TYPE: PR
 ; ORGANISM: Mus sp.
 ; FEATURE:
 ; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) B4 amino
 ; OTHER INFORMATION: acid sequence
 US-10-124-598-2

Query Match 70.9%; Score 3151.5; DB 12; Length 843;
 Best Local Similarity 68.9%; Pred. No. 2e-257;
 Matches 581; Conservative 113; Mismatches 142; Indels 7; Gaps 4;

QY 1 MGPRKATCSLFFLLWLAEP---AENSDFYLPDYLGLGSLSHANKGIHNLFIQVP 57
 Db 1 MGPRKATLHLFLHLHLHPKPVMLVGSDFHLAGDYLGLFTLHAAVKSYSHLSIQVP 60
 QY 58 MCKEYKVGIVGYNLMQAMRRAVEEINNDSSLLPGVLLGLEYIVDVCYSNNVQVLYELAH 117
 Db 61 KCNEYKNKVGYNLMQAMRRAVEEINNCSSLLPGVLLGLEYIVDVCYSNNVQVLYELAH 120
 QY 118 EDNLPLQEDYSNYSKRVAVAVIGDNSESVTANFSLFLPQITTSALSDERDVRF 177
 Db 121 IDDELPLIKYQYRPPVAVVIGDNSESAITVSNILSYLPQVYTSATLTKLQDRKF 180
 QY 178 PALRTPTSDHVEAVQVQLHFRMNIIVLVSSDTYGRNGQLGERVARR-DICIAF 236
 Db 181 PALRTPTSDHVEAVQVQLHFRMNIIVLVSSDTYGRNGQLGERVARR-DICIAF 240
 QY 237 QETPLTLPQNMNTSEERQRLVTIVDKLQOSTARVVVFSFDLTLYHFEVEVLKQNTGA 296
 Db 241 QEVLPPEPNOAVPEBODLDNLTKLRISARVVVIFSEELSLHNFREVLHMTGTF 300
 QY 297 VMTASEMAIDPVYHNLTELGHTGFTIGITQSPRIGESFERMGQAGPPISRTSOS 356
 Db 301 VMTASEMAIDPVYHNLTELGHTGFTIGITQSPRIGESFERMGQAGPPISRTSOS 360
 QY 357 YTCNQEDCNLNTLSFNTILRLSGERVVSVSAVAVAAHLSLGCSTCKRVVY 416
 Db 361 YTCNQEDCNLNTLSFNTILRLSGERVVSVSAVAVAAHLSLGCSTCKRVVY 420
 QY 417 PQOLLEIKVNTFLDHOIFPDQGVALHLETIVOMQDRSONPEOSVASTYPLQROK 476
 Db 421 PQOLLEIKVNTFLDHOIFPDQGVALHLETIVOMQDRSONPEOSVASTYPLQROK 480
 QY 477 NIDISHTVNTNTPMSCKRQSGOKKPVGIHVCCEPICDLPCTFLNHTDEVEQC 536
 Db 481 YISVSWSTYNTNTPMSCKRQSGOKKPVGIHVCCEPICDLPCTFLNHTDEVEQC 540
 QY 537 ACPPNEMSYOSETSCFKROLVFLFEMHAPITAVALLAALFSLTALIVFMRFOPIV 596
 Db 541 SCPSMSSTYKNNIACFKRLAFLEMHVPTIYVITIALAGFISTALILLFMRFOPIV 600
 QY 597 RSAGGPCFMTLLLVAVYVVPVYVPGVSTCLCROALFPLCFTICISIAVRSQIV 656
 Db 601 RSAGGPCFMTLLLVAVYVVPVYVPGVSTCLCROALFPLCFTICISIAVRSQIV 660
 QY 657 CAERMAERPRAYSIVYRGPIYSAFTIVLVKAVIVIGMLARPQSH--RIDDDPKI 714
 Db 661 CVKMAERPRAYSIVYRGPIYSAFTIVLVKAVIVIGMLARPQSH--RIDDDPKI 719
 QY 715 TIYSCNPNYSNLLENSTLULLSVGFSRAYMCKELPTVYNNEKFTTISMTFFTSYS 774
 Db 720 TIYSCNPNYSNLLENSTLULLSVGFSRAYMCKELPTVYNNEKFTTISMTFFTSYS 779

QY 775 ICTFMSAYSGVLTIVDLVTLNLAISLGEFGPKCMILFFPERNTPAVFSMIQGT 834
 Db 780 ICTFMSAYSGVLTIVDLVTLNLAISLGEFGPKCMILFFPERNTPAVFSMIQGT 839
 QY 835 MRR 837
 Db 840 MRR 842

RESULT 12

US-09-361-652-1
 ; Sequence 1, Application US/09361652
 ; Publication No. US20030036530A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zuker, Charles S.
 ; APPLICANT: Adler, Jon Elliot
 ; APPLICANT: Lindemeyer, Juergen
 ; APPLICANT: Ryba, Nick
 ; APPLICANT: Hoon, Mark
 ; TITLE OF INVENTION: The Regents of the University of California
 ; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
 ; FILE REFERENCE: 02307E-08610US
 ; CURRENT APPLICATION NUMBER: US/09/361,652
 ; CURRENT FILING DATE: 1999-07-27
 ; EARLIER APPLICATION NUMBER: US 60/094,465
 ; EARLIER FILING DATE: 1998-07-28
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 840
 ; TYPE: PR
 ; ORGANISM: Rattus sp.
 ; FEATURE:
 ; OTHER INFORMATION: rat G-protein coupled receptor B3 (GPCR-B3)
 US-09-361-652-1

Query Match 32.7%; Score 1454; DB 9; Length 840;
 Best Local Similarity 39.5%; Pred. No. 4.7e-114;
 Matches 325; Conservative 133; Mismatches 346; Indels 28; Gaps 11;

QY 27 FYLDGVDLGLGSLSHANKGIHNLFIQVPMCKEYKVGIVGYNLMQAMRRAVEEINND 85
 Db 31 FYLDGVDLGLGSLSHANKGIHNLFIQVPMCKEYKVGIVGYNLMQAMRRAVEEINND 88
 QY 86 SSLPLGLGLEYIVDVCYSNNVQVLYELAH--DNILPQEDYSNYSKRVAVAVIGDPS 144
 Db 89 SALLPNTTIGLEYIVDVCYSNNVQVLYELAH--DNILPQEDYSNYSKRVAVAVIGDPS 148
 QY 145 ESWVTANFSLFLPQITTSALSDERDVRPALIRTPSDHVEAVQVQLHFRMNI 204
 Db 149 DHAVTAAALGPFLMPLVSTSEASVYSAKRKPFSLRVPDSRHQVEAVQVQLHFRMNI 208
 QY 205 WIIYVSSDPTYGRNGQLGERVARRDICAFOETPLTLPQNMNTSEERQRLVTIVDKL 264
 Db 209 WIIYVSSDPTYGRNGQLGERVARRDICAFOETPLTLPQNMNTSEERQRLVTIVDKL 262
 QY 265 OOSTARVAVVESPDLTLYHFEVNTLRFONFTGAVYIASESWAIDPVYHNLTELGHTGFTIG 324
 Db 263 AQARTTYVVFVSNHRLARVFRSVYLANLTKGVAVASEDAISYITTSVIGIGIVLG 322
 QY 325 ITTQSVIPGFSEFERMGQAGPPISRTSOSYTC--NQCDCNCLNNTLSFNTILRLSGE 382
 Db 323 VAVQQRVPGLKEEESYVAVYAAAPACBEGSCSTNOICRCHFTTNNMPTLGAFSM 382
 QY 383 RVYYSVSAVAVAAHLSLGCSTCKRVVYVPMOLLEIKVNTFLDHOIFPDQ 442
 Db 383 SAARYVEAVYVAAHLSLGCSTCKRVVYVPMOLLEIKVNTFLDHOIFPDQ 442
 QY 443 DVALLHETIVOMQDRSONPEOSV--ASTYPLQROKRIQDIISHTVNTNTPMSCKRQ 500
 Db 443 DVLGYYDIAMDNNGPEWTEIIGSASLSPVHLDI--NKRRIQWGRKNQVAVSCTTDL 501

```

QY 501 SGKKKKPVGHVCCFECIDLCPLGTFELNHTDEBYCCQACPNMENSYSGETSCGRQVLVE 560
Db 502 AGHHRRVVGSHHCCFECVPCPCAGTFELNMSB-LHLCQPCGTEEMAPKSTTCCFPRTVEFLA 560
QY 561 WHEAPTAVALALLALGFLSTALIV-----IFMRHPOFPIVSAGBPOMCFMILTILVA 614
Db 561 WHE--PISLVILIA-----NILLILLVGTAGLEFMHHTHPVRSAGGRLLCFMIGSLVAG 614
QY 615 YMVVVVYGVPPKVSCTCLRQALPFLCFTICISIAVBSFOIVCAFKAASFRPRAYSQWVR 674
Db 615 SCSEFSPFGEPTVACLRQPLFSLGPAIFLSCLTITSFQVLIIFKSTKVPPIFYKRWQAQ 674
QY 675 YQGPVVSMAFTVLMKVLIVIGMLARPQSHRDPDDPKITIVSCNPNRYNSLFTNTSLD 734
Db 675 NHGAGLEFIVTSYVHLILCLMVLMMTPRPPIREXQRPPLVILECTEVNSVGFLLATPHN 734
QY 735 LLLSVVGSFRAYMGKELPNTNNAKFTTLSMTFFYFTSSVSLCJFMSAYSGVLATYDDL 794
Db 735 ILLISIFVCVSLKELEPENENAKCVTFSLLLNFVSWIAFTWASIVQSSYLPFAVNVA 794
QY 795 TVLNLALATISLGGPGKCMILFVBERNMPAFENMIGQYMR 836
Db 795 GLTTLSSGFSGLFPLKCVILLCRELNLNTEFQASIDQYTRR 836

```

```

RESULT 13
US-09-927-315-1
:
: Sequence 1, Application US/09927315
: Publication No. US20030040045A1
: GENERAL INFORMATION:
:
: APPLICANT: Zuker, Charles S.
: APPLICANT: Ryba, Nicholas J.P.
: APPLICANT: Nelson, Greg
: APPLICANT: Hoon, Mark A.
: APPLICANT: Chandrasekar, Jayaram
: APPLICANT: Zhang, Yifeng
: APPLICANT: The Regents of the University of California
: APPLICANT: The Government of the United States of America
: APPLICANT: as represented by the Secretary of the
: TITLE OF INVENTION: Department of Health and Human Services
: FILE REFERENCE: 02307E-120110US
: CURRENT APPLICATION NUMBER: US/09/927, 315
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: US 60/302, 898
: PRIOR FILING DATE: 2001-07-03
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO 1
: LENGTH: 840
: TYPE: prt
: ORGANISM: Rattus sp.
: FEATURE:
:
: OTHER INFORMATION: rat TrIR sweet taste receptor
:
: US-09-927-315-1

```

Query Match	32.7%	Score 1454;	DB 9;	Length 840;
Best Local Similarly	39.5%;	Pred. No. 4.7e-114;		
Matches 325; Conservative	123;	Mismatches 346;	Indels 28;	Gaps 11;

[illegible][illegible]

```

RESULT 14
US-09-361-652-2
: Sequence 2, Application US/09361652
: Publication No. US20030036630A1
: GENERAL INFORMATION:
: APPLICANT: Zuker, Charles S.
: APPLICANT: Adler, Jon Elliott
: APPLICANT: Lindemeier, Uergeren
: APPLICANT: Ryba, Nick
: APPLICANT: Hoon, Mark
: APPLICANT: The Regents of the University of California
: TITLE OF INVENTION: Nucleic Acids Encoding a G-protein Coupled Receptor
: TITLE OF INVENTION: Involved in Sensory Transduction
: FILE REFERENCE: 02307E-088610US
: CURRENT APPLICATION NUMBER: US/09/361,652
: CURRENT FILING DATE: 1999-07-27
: EARLIER APPLICATION NUMBER: US 60/094,465
: EARLIER FILING DATE: 1998-07-28
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2
: LENGTH: 842
: TYPE: prt
: ORGANISM: Mus sp.
: FEATURE:
: OTHER INFORMATION: mouse G-protein coupled receptor B3 (GPCR-B3)
US-09-361-652-2

```

Query Match 32.38; Score 1434; DB 9; Length 842;

Best Local Similarity 39.6%; Pred. No. 2.3e-112;
Matches 327; Conservative 123; Mismatches 340; Indels 36; Gaps 15;

```

27 FYLPGDYLGGFSLHNMKGIVLNFLOVPMC-KEYEVKIGVNIQMAFEAVEEIND 85
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
33 FSLPGDVLGGFSLHNMKGIVLNFLOVPMC-KEYEVKIGVNIQMAFEAVEEIND 90
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
86 SLLPGVLLGGEIVDVCYISNNVQVLYFLAHE-DNLLPFOEDYSNTISRVAVIGPDS 144
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91 TALLPNTTIGELYDVCSSESNVATLRVPAQGTGHEMQRLRNSSKVALIGDNT 150
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
145 ESWTVANFSLFLPQITYSAISDELBDKVRFPALLRTTPSADHVAHYAMQMLHRWN 204
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 DHAVTYALLSPFLMPVLYSEASSVILSGKRKFPFLRTIPSDRYQVEVYIRLLQSGWV 210
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
205 WIIVVSDTYGRNGQLGERVARRDICIAFOETPLTQPNOMTSEERQRLTYIDKL 264
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
211 WISLVSGYDGGOLGVALLELATPRGICVAFKDVPL---SAGADPRMORML---RL 264
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265 QOSTARVYVFPDLTYHFEVNEVLRONFTGAVWIAESMAIDVLLNLTLELHGLFTLG 324
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265 ARARTYVYVFSNRHLGAVFRRSVVLANLGKVMIASEDMAISTYITNVPGIOGIVLG 324
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
325 ITTOSVPIPGSEFERENGPOA--GPPPLSRTSOSTYCNQEDCNLANFLSNTILRSGE 382
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
325 VALQROVPGLKEFEESYVQAVMGAPRTCPREGSNGCTNOJCREFHAFTTMMPELAGFSM 384
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
383 RVVSVYSAYAVAHALSHLSDGCDSTCKRVVYPMOLLEIKVNFLLDHOJFFPDG 442
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
385 SAATVYEAAYAAVAGHGLHGLGCTSGTCARGPVYPMOLLQOITYVNFLLHKKYAFDDKG 444
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
443 DVALHLEIVOMQDRSONPQSV--ASYYPLOQRLKNIODISWHTVNTTIPMSCSKRCQ 500
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
445 DPLGYDIIDAMDNGPMTPEVIGSASLSPVHLDI-NKTKIQMHGKNNQVAVSVCTRCL 503
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
501 SGOKKRVGIVHCCFECIDCLPGFTLHTEDEYECQACPNNEYSOSETSCFKQVLFLE 560
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
504 EGHRLWVGHSHCCFECMPCEAGTFELN-TSELHQCQCGTEVMAPESSACFSRTVEFLG 562
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
561 WHEAPTVALLALGFLTLALIVF-----WRHFQPIVYSAGGPMCFMLTLLV 613
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
563 WHE--PISIVLAA-----NTLLLLLLIGTAGLFAFR-LHTPVVASAGRCFLMLGSLVA 615
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
614 AYVAVPYVGPVKVSTLCROALFPLCTTICISIAVRSFOICAFKMASFRPAYSYW 673
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
616 GSCSLYGFPGKPYVACILRQPLFSLGFAIFLSCLTIRSOQVLIIFRFSKVPFFYHTWA 675
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
674 RYOGPIVSMAPFIVLKAVIVIGMLARPOSHPTDPO--DPKITVYSCNPNTNSILFN 730
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
676 QNHGAGI---FVIYSSVTHFLCLTWMAMTPTRETREYQRFPHVILIECTEVNSVGLVA 732
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
731 TSIDLILSVGFSFAYNGKELPTNYNEAKFTILSMFTFSSVSLCTFMAVSGVLYTV 790
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
733 FAHNIILISTFVCSYIGKELPENYNEAKCVTSLLHFAVSWIAFTWSSYIGOSYLPAY 792
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
791 DLAVTVANLALISIGFPGKCYMLLFTPERNTPAVENSMDIOGYMR 836
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
793 NVLAGLTLTSGGFSGYFLPKCYVILCRPELNTNTEHFOASIODYTRR 838
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 15
US-09-927-315-2

; Sequence 2, Application US/09927315
; Publication No. US20030040045A1

; GENERAL INFORMATION:

; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nicholas J.P.

; APPLICANT: Nelson, Greg
; APPLICANT: Hoon, Mark A.

; APPLICANT: Chandrasekar, Jayaram
; APPLICANT: Zhang, Yifeng

; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America

```

; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
; FILE REFERENCE: 02307E-120110US
; CURRENT APPLICATION NUMBER: US/09/927, 315
; PRIOR FILING DATE: 2001-08-10
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: mouse Tr1 sweet taste receptor
US-09-927-315-2

```

Query Match

Best Local Similarity 32.3%; Score 1434; DB 9; Length 842;

Matches 327; Conservative 123; Mismatches 340; Indels 36; Gaps 15;

```

27 FYLPGDYLGGFSLHNMKGIVLNFLOVPMC-KEYEVKIGVNIQMAFEAVEEIND 85
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
33 FSLPGDVLGGFSLHNMKGIVLNFLOVPMC-KEYEVKIGVNIQMAFEAVEEIND 90
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
86 SLLPGVLLGGEIVDVCYISNNVQVLYFLAHE-DNLLPFOEDYSNTISRVAVIGPDS 144
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
91 TALLPNTTIGELYDVCSSESNVATLRVPAQGTGHEMQRLRNSSKVALIGDNT 150
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
145 ESWTVANFSLFLPQITYSAISDELBDKVRFPALLRTTPSADHVAHYAMQMLHRWN 204
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 DHAVTYALLSPFLMPVLYSEASSVILSGKRKFPFLRTIPSDRYQVEVYIRLLQSGWV 210
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
205 WIIVVSDTYGRNGQLGERVARRDICIAFOETPLTQPNOMTSEERQRLTYIDKL 264
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
211 WISLVSGYDGGOLGVALLELATPRGICVAFKDVPL---SAGADPRMORML---RL 264
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265 QOSTARVYVFPDLTYHFEVNEVLRONFTGAVWIAESMAIDVLLNLTLELHGLFTLG 324
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265 ARARTYVYVFSNRHLGAVFRRSVVLANLGKVMIASEDMAISTYITNVPGIOGIVLG 324
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
325 ITTOSVPIPGSEFERENGPOA--GPPPLSRTSOSTYCNQEDCNLANFLSNTILRSGE 382
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
325 VALQROVPGLKEFEESYVQAVMGAPRTCPREGSNGCTNOJCREFHAFTTMMPELAGFSM 384
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
383 RVVSVYSAYAVAHALSHLSDGCDSTCKRVVYPMOLLEIKVNFLLDHOJFFPDG 442
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265 ARARTYVYVFSNRHLGAVFRRSVVLANLGKVMIASEDMAISTYITNVPGIOGIVLG 324
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
325 ITTOSVPIPGSEFERENGPOA--GPPPLSRTSOSTYCNQEDCNLANFLSNTILRSGE 382
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
325 VALQROVPGLKEFEESYVQAVMGAPRTCPREGSNGCTNOJCREFHAFTTMMPELAGFSM 384
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
383 RVVSVYSAYAVAHALSHLSDGCDSTCKRVVYPMOLLEIKVNFLLDHOJFFPDG 442
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
385 SAATVYEAAYAAVAGHGLHGLGCTSGTCARGPVYPMOLLQOITYVNFLLHKKYAFDDKG 444
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
443 DVALHLEIVOMQDRSONPQSV--ASYYPLOQRLKNIODISWHTVNTTIPMSCSKRCQ 500
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
445 DPLGYDIIDAMDNGPMTPEVIGSASLSPVHLDI-NKTKIQMHGKNNQVAVSVCTRCL 503
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
501 SGOKKRVGIVHCCFECIDCLPGFTLHTEDEYECQACPNNEYSOSETSCFKQVLFLE 560
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
504 EGHRLWVGHSHCCFECMPCEAGTFELN-TSELHQCQCGTEVMAPESSACFSRTVEFLG 562
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
561 WHEAPTVALLALGFLTLALIVF-----WRHFQPIVYSAGGPMCFMLTLLV 613
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
563 WHE--PISIVLAA-----NTLLLLLLIGTAGLFAFR-LHTPVVASAGRCFLMLGSLVA 615
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
614 AYVAVPYVGPVKVSTLCROALFPLCTTICISIAVRSFOICAFKMASFRPAYSYW 673
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
616 GSCSLYGFPGKPYVACILRQPLFSLGFAIFLSCLTIRSOQVLIIFRFSKVPFFYHTWA 675
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
674 RYOGPIVSMAPFIVLKAVIVIGMLARPOSHPTDPO--DPKITVYSCNPNTNSILFN 730
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
676 QNHGAGI---FVIYSSVTHFLCLTWMAMTPTRETREYQRFPHVILIECTEVNSVGLVA 732
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
731 TSIDLILSVGFSFAYNGKELPTNYNEAKFTILSMFTFSSVSLCTFMAVSGVLYTV 790
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
733 FAHNIILISTFVCSYIGKELPENYNEAKCVTSLLHFAVSWIAFTWSSYIGOSYLPAY 792
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Sun May 25 15:06:02 2003

us-09-927-315-9.rapb

Page 11

OY 791 DLIVVNLIALSLGFFGPCYMILEYPERTPAYFNSMIGYTMR 836
:: : | ||||| : | : : | |
Db 793 NYLAGLATLSGGSGVFLLPCRYILCRDELNTEHFGQSIDYTRR 838

Search completed: May 19, 2003, 10:04:41
Job time : 43.1769 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 11:21:44 ; Search time 5176.99 seconds

(without alignments)
11299.347 Million cell updates/sec

Title: US-09-927-315-12

Perfect score: 2010

Sequence: 1 ataccctacagcgcacatcag.....acacccatgagggagactag 2010

Scoring table:

IDENTITY:NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:.*
1: gb_da.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rnd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2010	100.0	2010	6	AR208617 Sequence
2	1232.8	61.3	2993	6	AR208615 Sequence
3	1230	61.2	2529	10	AF127390 Ratius no
4	1220.6	60.7	3060	10	AF337041 Mus muscu
5	1212.6	60.3	2532	10	AR208616 Sequence
6	1212.6	60.3	2532	10	AY032623 Mus muscu
7	879.8	43.8	930	2	AF45814956 Homo sapi
8	878.2	43.7	209676	2	AL831755 Homo sapi
9	774	38.5	774	9	AF45814953 Homo sapi
10	605	30.1	206074	2	AL831790 Homo sapi
11	474.6	23.6	154891	2	AC115397 Ratius no
12	392.8	19.5	2520	10	AF127389 Ratius no
13	386.4	19.2	2292	6	AX282894 Sequence
14	386.4	19.2	2526	6	AX282892 Sequence
15	386.4	19.2	2951	6	AX282900 Sequence
16	385.2	19.2	2759	6	AX354027 Sequence
17	384.8	19.1	2516	9	AF387618 Homo sapi
18	367.2	18.3	2526	10	AF301162 Mus muscu
19	359.2	17.9	2529	10	AY032622 Mus muscu
20	359.2	17.9	2892	10	AF337040 Mus muscu
21	357.6	17.8	2526	10	AF301161 Mus muscu
22	258	12.8	697	6	AX320494 Sequence
23	218.6	10.9	3458	6	AX417279 Sequence
24	217	10.8	2559	6	AX239649 Sequence
25	217	10.8	2559	6	AX298370 Sequence
26	217	10.8	3489	6	AX239647 Sequence
27	190.4	9.5	10578	9	AB065618 Homo sapi
28	190.4	9.5	102539	9	AL591866 Human DNA
29	190.4	9.5	191481	2	AC108123 Homo sapi
30	187	9.3	4116	5	AF406649 Squalus a
31	187	9.3	4134	6	AR182343 Sequence
32	187	9.3	4134	6	AX417770 Sequence
33	185.4	9.3	4134	6	AX449101 Sequence
34	185.4	9.2	1383	9	AF387619 Homo sapi
35	184.2	9.2	1101	6	AX282896 Sequence
36	184.2	9.2	93897	2	AC129831 Ratius no
37	182.8	9.1	210	9	AF45814954 Homo sapi
38	180.2	9.0	2729	6	AX354026 Sequence
39	172.6	8.6	5249	4	S67307 Ca(2+)-sens
40	172.6	8.6	5275	6	AR012622 Sequence
41	172.6	8.6	5275	6	AR028465 Sequence
42	172.6	8.6	5275	6	AR078217 Sequence
43	172.6	8.6	5275	6	AR177781 Sequence
44	172.6	8.6	5275	6	IR75051 Sequence I
45	168.6	8.4	173724	2	AC026766 Homo sapi

ALIGNMENTS

RESULT 1
AR208617
LOCUS AR208617 2010 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 8 from patent US 6383778.
ACCESSION AR208617
VERSION AR208617.1 GI:21509818
KEYWORDS
SOURCE
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 2010)
AUTHORS Zuker,C.S., Adler,J.E. and Lindemeier,J
TITLE Nucleic acids encoding a G-protein coupled receptor involved in
sensory transduction
JOURNAL Patent: US 6383778-A 8 07-MAY-2002;

```

FEATURES { Location/Qualifiers
    source             1..2010
                        /organism="unknown"
BASE COUNT      387 a      691 c      515 g      417 t
ORIGIN
Query Match      100.0%; Score 2010; DB 6; Length 2010;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACCTACAGCGCCATCAGAGATGAGTGGAGACAAAGGTCCTCCGGCTTGGCTG 60
DB 1 ATGACCTACAGCGCCATCAGAGATGAGTGGAGACAAAGGTCCTCCGGCTTGGCTG 60
QY 61 CGTACCAACCCAGCCGCGACCAACACAGTCGAGGCCATGGTGAAGTATGCTGCACTTC 120
DB 61 CGTACCAACCCAGCCGCGACCAACACAGTCGAGGCCATGGTGAAGTATGCTGCACTTC 120
QY 121 CGCTGGAACTGATCATTTGCTGTGTGAGACAGCAGACACTATGGCCGCGACATGGCCAG 180
DB 121 CGCTGGAACTGATCATTTGCTGTGTGAGACAGCAGACACTATGGCCGCGACATGGCCAG 180
QY 181 CTGCTTGGCAGCGCCGCGCGCGCGAGATGTCATGCGCTTCCAGAGAGCGTGGCC 240
DB 181 CTGCTTGGCAGCGCCGCGCGCGCGAGATGTCATGCGCTTCCAGAGAGCGTGGCC 240
QY 241 ACACTGAGCCCAACCAACCAATGATGCTCAGAGAGCGCCAGCGCTGTGATGATGTG 300
DB 241 ACACTGAGCCCAACCAACCAATGATGCTCAGAGAGCGCCAGCGCTGTGATGATGTG 300
QY 301 GACAACTGACAGCAGACAGACAGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
DB 301 GACAACTGACAGCAGACAGACAGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
QY 361 TACCACTTCTTCAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
DB 361 TACCACTTCTTCAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 421 GAGTCTGGGCGCATGACCGCGGTCTGTGACAACTCAGGAGCTGGGCGCATTTGGGCGAC 480
DB 421 GAGTCTGGGCGCATGACCGCGGTCTGTGACAACTCAGGAGCTGGGCGCATTTGGGCGAC 480
QY 481 TTCTGTGGGCGCATGACCGCGGTCTGTGACAACTCAGGAGCTGGGCGCATTTGGGCGAC 540
DB 481 TTCTGTGGGCGCATGACCGCGGTCTGTGACAACTCAGGAGCTGGGCGCATTTGGGCGAC 540
QY 541 GGGCGACAGGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 GGGCGACAGGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 GAGTGGACAACTGCTGGAAGCGCCACTTCTCTTCAACCACTTCTCAGGCTCTGGG 660
DB 601 GAGTGGACAACTGCTGGAAGCGCCACTTCTCTTCAACCACTTCTCAGGCTCTGGG 660
QY 661 GAGGCTGTGCTACAGCGTACTGTGCGGTCTATGCTGTGGCGCATTTGGCGCAAGC 720
DB 661 GAGGCTGTGCTACAGCGTACTGTGCGGTCTATGCTGTGGCGCATTTGGCGCAAGC 720
QY 721 CTCTCTGGCTGTGCAAAACACCTTGACCAAGAGGCTGTCTTACCTCTGGACACTGCTT 780
DB 721 CTCTCTGGCTGTGCAAAACACCTTGACCAAGAGGCTGTCTTACCTCTGGACACTGCTT 780
QY 781 GAGGAGATCTGGAAGGTCAACTTCTCTCTGAGACCAACCAATCTTCTGACCCGCAA 840
DB 781 GAGGAGATCTGGAAGGTCAACTTCTCTCTGAGACCAACCAATCTTCTGACCCGCAA 840
QY 841 GGGGAGCTGGCTGTGCACTTGTGAGATGTGCAAGTGGCAATGGGCGGAGCGCAATGCC 900
DB 841 GGGGAGCTGGCTGTGCACTTGTGAGATGTGCAAGTGGCAATGGGCGGAGCGCAATGCC 900
QY 901 TTCCAGAGCGTCTGCTCTACTACCCCTGACAGAGCAGCTGAAGAATCAAGACATCT 960
DB 901 TTCCAGAGCGTCTGCTCTACTACCCCTGACAGAGCAGCTGAAGAATCAAGACATCT 960

```

```

QY 961 CTGCACACCGTCAACACAGATCCATGTGCTTCAAGAGGTCCAGTCAGG 1020
DB 961 CTGCACACCGTCAACACAGATCCATGTGCTTCAAGAGGTCCAGTCAGG 1020
QY 1021 CAAAAGAGAGCGCTGTGGCATTCACAGTGTGCTCTTCAAGTATGCTGCTTCC 1080
DB 1021 CAAAAGAGAGCGCTGTGGCATTCACAGTGTGCTCTTCAAGTATGCTGCTTCC 1080
QY 1081 GGCACCTTCTTCAACCACTGATGATGCGGATTAACAGAGTGTCTCAGAGTGAAC 1140
DB 1081 GGCACCTTCTTCAACCACTGATGATGCGGATTAACAGAGTGTCTCAGAGTGAAC 1140
QY 1141 TCCGTCTTCAAGCGGACAGTGTCTTCTGTAATGGCATGAGGACCCACATCGCTGTG 1200
DB 1141 TCCGTCTTCAAGCGGACAGTGTCTTCTGTAATGGCATGAGGACCCACATCGCTGTG 1200
QY 1201 GCGCTGTGGCGCGCTTGGGCTTCTTACAGACCCCTGGCCATCTCTGTGATTTCTGAGG 1260
DB 1201 GCGCTGTGGCGCGCTTGGGCTTCTTACAGACCCCTGGCCATCTCTGTGATTTCTGAGG 1260
QY 1261 CACTTCCAGACACCCATATGTTGGTGGGCGCGCATGTGCTTCTGATGCTGACA 1320
DB 1261 CACTTCCAGACACCCATATGTTGGTGGGCGCGCATGTGCTTCTGATGCTGACA 1320
QY 1321 CTGCTGTGGTGGCATATGATGTGTCCCGGTGTACGTGGGCGCGCAAGTCTCCACC 1380
DB 1321 CTGCTGTGGTGGCATATGATGTGTCCCGGTGTACGTGGGCGCGCAAGTCTCCACC 1380
QY 1381 TGCCTTGGCGCGCGCGCTTCTTCCCTCTGCTGTCAATTTGCAATCTCTGTATGCC 1440
DB 1381 TGCCTTGGCGCGCGCGCTTCTTCCCTCTGCTGTCAATTTGCAATCTCTGTATGCC 1440
QY 1441 GTGCGTCTTCCAGATGCTGTGCGCTTCAAGATGAGCGCGCTTCCAGCGGCTAC 1500
DB 1441 GTGCGTCTTCCAGATGCTGTGCGCTTCAAGATGAGCGCGCTTCCAGCGGCTAC 1500
QY 1501 AGCTACTGGGTCCGCTACAGAGGCGCGCTACGCTCTATGAGCATTTATACGGTACTCAA 1560
DB 1501 AGCTACTGGGTCCGCTACAGAGGCGCGCTACGCTCTATGAGCATTTATACGGTACTCAA 1560
QY 1561 ATGGTCAATGGTAATTTGATGCTGTGGAGCGGCGCTACGCGCGCGCTACGCGCGCG 1620
DB 1561 ATGGTCAATGGTAATTTGATGCTGTGGAGCGGCGCTACGCGCGCGCTACGCGCGCG 1620
QY 1621 GATGACCCCAAGATCACAATTTCTCTGTATACCCCACTACCGCAACAGCTGTGTT 1680
DB 1621 GATGACCCCAAGATCACAATTTCTCTGTATACCCCACTACCGCAACAGCTGTGTT 1680
QY 1681 AACACAGCGCTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1740
DB 1681 AACACAGCGCTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1740
QY 1741 GAGTGGCGCAACCACTCAACAGAGGCGCAAGTATATACCCCTCAGCATGACCTTATTT 1800
DB 1741 GAGTGGCGCAACCACTCAACAGAGGCGCAAGTATATACCCCTCAGCATGACCTTATTT 1800
QY 1801 ACCTCATCGCTTCTCTCTGACCTTATGCTGTGCTACAGCGGCGGTGTGTACCATC 1860
DB 1801 ACCTCATCGCTTCTCTCTGACCTTATGCTGTGCTACAGCGGCGGTGTGTACCATC 1860
QY 1861 GTGAGCTCTTGTGCTACTGTGTCAACCTCTGTGGCATCAGCTGTGGCTACTTGGGCGCC 1920
DB 1861 GTGAGCTCTTGTGCTACTGTGTCAACCTCTGTGGCATCAGCTGTGGCTACTTGGGCGCC 1920
QY 1921 AAGTGTACATGATTCCTTCTTACCCGAGAGCGCAACGCGCGCTTACTTCAACAGATG 1980
DB 1921 AAGTGTACATGATTCCTTCTTACCCGAGAGCGCAACGCGCGCTTACTTCAACAGATG 1980
QY 1981 ATCCAGGCTACACCATGAGAGGAGCTAG 2010
DB 1981 ATCCAGGCTACACCATGAGAGGAGCTAG 2010

```


RESULT 2
LOCUS AR208615 2993 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 3 from patent US 6383778.
ACCESSION AR208615
VERSION AR208615.1 GI:21509815
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2993)
AUTHORS Zuker,C.S., Adler,J.E. and Lindemeyer,J.
TITLE Nucleic acids encoding a G-protein coupled receptor involved in
sensory transduction
JOURNAL Patent: US 6383778-A 3 07-MAY-2002;
FEATURES
source location/Qualifiers
1..2993
BASE COUNT 619 a 962 c 729 g 683 t
ORIGIN
Query Match 61.3%; Score 1232.8; DB 6; Length 2993;
Best Local Similarity 76.9%; Pred. No. 9.6e-194;
Matches 1568; Conservative 0; Mismatches 442; Indels 30; Gaps 4;
QY 1 ATCACTACAGGCGCATAGCATAGTGGAGACAAGAGTGGCTTCCGGCTTGGT 60
DB 505 ATCACTACAGGCGCATAGCATAGTGGAGACAAGAGTGGCTTCCGATGCTA 564
QY 61 CGTACCAACACCGGCGGACACGACGTCGAGGCGCATGTCAGTGTGTCATTC 120
DB 565 CGCAAGTGGCCAGGCGGACCCACACATCGAGGCGCATGTCAGTGTGTCATTC 624
QY 121 CGCTGAAGTGGATCATGTCGTGTGAGACAGCAGCACTATGCGCGGACAAATG 180
DB 625 CAATGAACTGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 684
QY 181 CTGCTTGGAGCGGTGGCC---CGGCGGCAATCTGTCATGCTTCCAGAGACGTG 237
DB 665 CTGTGAGCCAGCGTCTACCAAAAGAGCGACATCTGATGCTTCCAGAGGTTCG 744
QY 238 CCCACATGACGCCCAACAGACATGATGATGAGAGAGCGGCGGCGGACAAATG 297
DB 745 CCCATACCTGATGTCAGGCGGATGATGATGATGATGATGATGATGATGATG 804
QY 298 GTGACAAAGTGCAGACAGACAGCGCGTGTGTGTGTGTGTGTGTGTGTGTGT 357
DB 805 CTGGACAAAGTGCAGACATCGCGCGCGGTGTGTGTGTGTGTGTGTGTGTGT 864
QY 358 CTGTACCACTTCTCATGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 417
DB 865 CTGTACAGTCTTCTCATGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 924
QY 418 TCCGAGTCTTGGGCGCATGACCGGCTTGTGCACAACTGACGAGAGTGGGCGT 477
DB 925 TCTGAGTCTTGGGCGCATGACCGGCTTGTGCATACCTGACGAGAGTGGGCGT 984
QY 478 ACCTTCTTGGGCGCATGACCGGCTTGTGCATACCTGACGAGAGTGGGCGT 537
DB 985 ACTTCTTGGGCGCATGACCGGCTTGTGCATACCTGACGAGAGTGGGCGT 1044
QY 538 TGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 597
DB 1045 CGCGGTGACAAAGCGGATGTCGCTGTACAGGACGACGCGGAGCGAGCTGCA 1104
QY 598 CAGGAGTGCAGACCTGCTGAAGCGCACTGCTTCAACACCACTTCAAGGCTCT 657
DB 1105 CAGGAGTGCAGGCGCTGTGAACACACAGCTTCAACACCACTTCAAGGCTCT 1164
QY 658 GGGAGGCTGT 717
DB 1165 GGGAGGCTGT 1224

QY 718 AGCTTCTGCGGT 777
DB 1225 AGCTTCTGCGGT 1284
QY 778 GTTGAAGATGTGAAGTCACTTCACTGTCGAGACCAATCTTCTTTCGACCG 837
DB 1285 CTCAGGAGATGTGACAGTCACTTCACTGTCGAGACCAATCTTCTTTCGACCA 1344
QY 838 CAAGGGAGT 897
DB 1345 CAAGGGAGATGTGACAGTCACTTCACTGTCGAGACCAATCTTCTTTCGACCA 1404
QY 898 CCTTTCAGAGGT 954
DB 1405 CCTTTCAGAGGT 1464
QY 955 ACATCTGTGACACCGTCAACACAGATTCCTATGTCCATGTGTTCAGAGTGCAG 1014
DB 1465 GTGTCTGT 1524
QY 1015 TCAGGGCAAAAGAGAGAGCGTGTGGCATTCAGTGTGTGTGTGTGTGTGTGTGT 1074
DB 1525 CCAGGGCAAAATGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1584
QY 1075 CTTCGGGCGACCTTCTCAACACACTGAA-----TGCCGCAAT 1113
DB 1585 ATGCGAGGACCTTACCTCAACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1644
QY 1114 AACGAGT 1173
DB 1645 TCCATGT 1704
QY 1174 TGGCATGAGGCAACCAATGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1233
DB 1705 TGGCAGGAAGT 1764
QY 1234 CTGGGCAATCTGT 1293
DB 1765 CTGGGCAATCTGT 1824
QY 1294 GGGCCCATGT 1353
DB 1825 GGGCCCATGT 1884
QY 1354 TACGT 1413
DB 1885 TATGT 1944
QY 1414 TTCACAATTTGATGTCGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1473
DB 1945 TTCTCATCTGT 2004
QY 1474 ATGGCAGCGGCTTCCACAGCGCTTACAGCTACTGAGTCCGCTTACAGAGGCGCTT 1533
DB 2005 ATGGCAGAGCGGCTTCCACAGCGCTTACAGCTACTGAGTCCGCTTACAGAGG 2064
QY 1534 TCTATGCGATTTATACAGGTACTCAAAATGTGTGTGTGTGTGTGTGTGTGTGT 1593
DB 2065 TTTGT 2124
QY 1594 CCGT 1650
DB 2125 ACCATCAACCGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2184
QY 1651 AACCCCACTACCGGACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1710
DB 2185 CACCTTACTACCGGACAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2244
QY 1711 GTGGTTTACGTTGT 1770
DB 2245 CTGGGTTTACGTTGT 2304

OY	1771	TTCATCACCCCTAGCATGACCTTTCATTATTCACCTATCGCCTGCCCTGGACCCTTAG	1830
OY	1771		
DB	2305	TTCATCACTCTAGCATGACCTTTCTTCCCTTCACTCTCCATCTCCCTTGACACTTTCATG	2364
OY	1831	TCTGCCTACAGCGGGGTGCTGTGTCAACATCTGTGACCTCTTGTGTCACGTGTCTAACCTC	1890
DB	2365	TCTGTGACAGAGAGCGGTGCTGTGTCAACATCAAGACCTCTGTGTCACGTGTCTAACCTTC	2424
OY	1891	CTGGCATCAGCGCTGGGCTACTTGGCCCCAAGTCTCTACATGATTCCTCTTACCCGGAG	1950
DB	2425	CTGGCATCAGCGGTGGGATCTTGTGGCCCCAAGTCTCTACATGATTCCTTTCACCCGGAG	2484
OY	1951	CGCAACACGCCCCCTACTTCTTAACAACATGATTCACGAGGCTACACCATGAGAGGACTAG	2010
DB	2485	CGCAACACCTAGCCTATTCTTCAATAGCATGATTCACGAGGCTACACCATGAGAGAAGCTAG	2544
RESULT 3			
LOCUS	AF127390	2529 bp	mRNA linear ROD 04-MAR-1999
DEFINITION	Rattus norvegicus putative taste receptor TR2 mRNA, partial cds.		
VERSION	AF127390		
KEYWORDS	AF127390.1 GI:4337087		
SOURCE	Rattus norvegicus.		
ORGANISM	Rattus norvegicus.		
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Mammalia; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. 1 (bases 1 to 2529) Hoon,M.A., Adler,E., Lindemeyer,J., Battey,J.F., Ryba,N.J.P. and Zucker,C.S. Putative mammalian taste receptors: a class of taste-specific GPCRs with distinct topographic selectivity Cell 96 (4), 541-551 (1999)		
JOURNAL	99159821		
MEDLINE	10052456		
PUBMED	2 (bases 1 to 2529) Hoon,M.A., Adler,E., Lindemeyer,J., Battey,J.F., Ryba,N.J.P. and Zucker,C.S. Direct Submission Submitted (10-FEB-1999) Taste- and Smell unit, NIDCR, 10 Center Drive MSC 1188, Bethesda, MD 20892-1188, USA		
AUTHORS	Hoon,M.A., Adler,E., Lindemeyer,J., Battey,J.F., Ryba,N.J.P. and Zucker,C.S.		
FEATURES	Location/Qualifiers 1..2529 /organism="Rattus norvegicus" /strain="Wistar" /db_xref="taxon:10116" /cell_type="circumvallate papilla" 1..>2529 /note="g protein-coupled receptor" /codon_start=1 /product="putative taste receptor TR2" /protein_id="A018070.1" /db_xref="GI:4337088" /translaton="MGPOARTLCILSLHLVLPKGLIVENSDFHLADVLGLFTLL HANYSLSHSYLQVPCNFEETMKVFLNGYNLMQAPRAVEEINNSILLPGVILGEEM DVCLSNHHIGLAEFLAODDLDPLINDGYMHVAIVGDNSSALITNYLTSEHF LIPIITYSAISDKLRDRHPMSRLTPSATHTHEFAAQOLAVHEROMNTVIVASDDY GRESHLSLSORLTKTSIDICIAFOEVLPSPESOVRSBEBOQNDNLIDLKLRISARVV VPEPSELISYEFHIEVRWNETGWPMWIASMAIDPVLMHTELRHGTGLGVTLQVRI SIPEQSQRVARBRDPGPVPWTMLRTQCNOCDACLNTRKSFNIIILSGERVVSV YSAYAAVAHALRHLLRGNCNRVCTROKYPMQLAREIHWNETLGNLFEPDOOSDOM LDDILOMDLSONPOSISIASYSPSTRKIYINNVSPTNNPYVPSCKSDQDM KKSGVLRHCCECLDCMRGTYLINLSABRFNCLSPGSMSKTDITCFORPTLEHN EVPITVAIALAAGFESTLALFLFMHFQDPMVASGGPCFMLVPLLALAFEMVAV YVGFPAVSCRCQAFTVCFSICLSITVNSQIVCFKARLPSAYSFWMTKHQB VVFPAFTIKVALVVGNMALATTINPLGRDPPDDPINILSKRPYENGLETFSMDL LTVLSGFEFAAGKEELPTNVEAKFTILSMFTSTSSISICTEMSVDGLVLTIMDLL VTVALNPALIGTGYGRCYMLLPEDENTSATSFMSNIQTMRKS"		
CDS			
BASE COUNT	497 a	826 c	623 g 503 t
ORIGIN			

Query Match	Similarity	61.2%	Score 1230:	DB 10:	Length 2529:
Best Local	Similarity	76.9%	Pred. No. 2.8e-193:		
Matches 1564:	Conservative	0:	Mismatches 440:	Indels 30:	Gaps
QY	1	ATCACCTACAGCGCCATCAGCGATGAGCTGACAGACAAAGGTGCGCTTCCGGCTTTGCTG	60		
Db	493	ATTCATATACAGCGCCATCTCCGAAAGCTGGCGGAGCAAGCGGACACTTCCCTAGACATGTA	552		
QY	61	CGTACACACCCAGGGCCGACACCAACGTTCAGAGCCCATGGTGCATGTATGCTGCATCTTC	120		
Db	553	CGCACAGTCCCGAGCGCCACCCACACATCAGAGGCCCATGTGCTGAGTATGTTACATCTTC	612		
QY	121	CCGCTGAACCTGATCATTTGCTGGTGTGACAGCAGCACACTATGCGCGGACACATGGCCAG	180		
Db	613	CAATGGAACATGGAATTTGTGTGTGCTGGTGTGACGACGACGATTTACGGCGCGGAGAAACAGCCAC	672		
QY	181	CTGCTTTGGCGAGACGGCTGGCC---CGGCGCGACATCTGATGTCCTTCCAGAGACGCTG	237		
Db	673	CTGTTGAGCGACACGCTCTGACAAAGACGACGACATCTTGATTTGCCCTCCAGGAGGTTTCTG	732		
QY	238	CCCACTACTCAGCCCAACACGAAACATGACGTGAGAGAGAGCGCGACGCGCTGGTGCACAT	297		
Db	733	CCCATATACCTGATATCCAGCGCATGATGAGAGTCCGAGAGACACAGACATATGGACACATAC	792		
QY	298	GTGACACAGCTGCAGACAGACACAGCGCGCTGCTGTGCTGCTTCTCGCCGACCTGACC	357		
Db	793	CTGGACAAAGCTCGCGGGGACCTCGCGGCGCGCTGTGTGTGTGCTTCGCGCCGACGCTGAC	852		
QY	358	CTGTACCACTTCTTCAATAGGTGGCGGCGCGCAAGAACTTACGAGCGCGCCCTGGTGAATCGCC	417		
Db	853	CTGTATAGCTTTTTCACGAGGTGCTCCGCTGGAACCTTACCGGTTTGTGTGTGATCGCC	912		
QY	418	TCCGAGTCTTGAGGACATCGACCCGGTCTCTGACAAACCTCAGGAGCTGGGCCACTTGGGC	477		
Db	913	TCTGTAGTCTTGAGGCTATCCAGCCAGTTCGTGCATTAACCTCAGGAGCTGGCGCACAGGGGT	972		
QY	478	ACCTTCCTGGGGATACCAATTCGAGAGCGGCCCATCCGGGCTTCAAGTAGTTCGCGGAG	537		
Db	973	ACTTTTCTGGGGGTCAACATCCAGAGGGGTGCATCTCCGTCTCAGTACGTATCGAGTGTG	1032		
QY	538	TGGGGCCACACAGGCTGGGGCCACCCCTCAGACAGGACCCAGGCAGAGCTATACCTGCAAC	597		
Db	1033	GCCCGAGACAAACGACAGGGGATATCCCGGCTTAACAGACCAACCTGGGAGCAGACTGCAAC	1092		
QY	598	CAGGAGTGCAGACATGCTGCTGAAGCCACCTTGTCTTAAACATTTCTCAGGCTCTCT	657		
Db	1093	CAGGACTGTGAAGCTGCTGTTGAACACCAACCAAGTCCCTTCAACAAATCCTTATCTTGTG	1152		
QY	658	GGGAGACGCTGCTGTACACAGCTGTACTCTGTGGGCTATGCTGTGGGCCATACCCCTGCAC	717		
Db	1153	GGGAGACGCGGTGTGTACAGCTGTACTCTGTGGGCTATGCTGTGGGCCATACCCCTGCAC	1212		
QY	718	AGCCTCTCGCTGTGTGACAAAAAGCACTTGACCAAGAGGGGTGTACCCCTGGGACGCTG	777		
Db	1213	AGACTCTCTCGCTGTGAACCGGGGTCCGCTGACCAAGCAAAAGGTGTACCCGCTGGAGCTTA	1272		
QY	778	CTTGAAGGATGTGGAAGGTCAACTTCACTTCCTGGAGACCAACAATCTTTCTGCACCGG	837		
Db	1273	CTCAGGGAATGTGGACCTCACTTCACTTCCTGGGTTACCGGGCTCTTCTTGTGACAA	1332		
QY	838	CAAGGGAGAGTGGCTCTGTGCATTTGAGATTTGCCAGTGGCAATGGAGCCGAGCCAGAA	897		
Db	1333	CAAGGGAGACATGCCGATGTCTTTGGACATCATCAAGTGGCAAGTGGGAGCTGAGCCAGAA	1392		
QY	898	CCCTTCACAGAGCTGCTCTCTACTAACCCTCTGCACGACGACAGCTGAAGACAT---CAAG	954		
Db	1393	CCCTTCACAAAGCATGCCCTCTACTTCTCCACACAGCAAGAGGCTTAACCTCACTTAACAT	1452		
QY	954	ACATCTGCACACAGCTGAACAACAACAGATCCCTATGTCCATGTGTTCCAAAGGGGCGAG	1014		
Db	1453	GTGTCTGTGTACACCCCAACAACAGGTCTCTGTCTCATGTGTGTCCAAAGAGCTGGCCAG	1512		
QY	1015	TCAGGGCAAAACAAAGAGCTGTGGGCAATCCACAGTCTGTGCTTGGTAGTGCATGCATCC	1074		

```

Db 1513 CCAGGGCAATGAAAGTGTGTGGGCTCCACCTTGTCTTGATGCTGTGATTTGT 1572
QY 1075 CTTCGGGCACTTCTTCACACACATGA-----TGCCGCAAT 1113
Db 1573 ATGCCAGGCACTTACCTCAACCGCTCCAGCATGATTAACTGTCTGTCTCCGCGGGT 1632
QY 1114 AAGGAGTGTCTTCCAGAGTGTGAGACCTCCGCTTCAAGGGGAGCTGTCTGTGAA 1173
Db 1633 TCCATGTGTCTTCAAGAACACATCACTGTCTTCCAGGGGCTTACCTTCTGTGAG 1692
QY 1174 TGGCATGAGGCAACCACTGCTGTGTGGCTTGTGGGCTTGTGGGCTTCTTCAAGC 1233
Db 1693 TGGCAGCAAGTGGCCACCATGCTGTGTGGCTTGTGGGCTTGTGGGCTTCTTCAAG 1752
QY 1234 CTGGCCATCTCTGTGTATTTGTGAGGACATTCAGACACCATATGTTGCTGGCTGG 1293
Db 1753 CTGGCCATCTCTGTGTATTTGTGAGGACATTCAGACACCATATGTTGCTGGCTGG 1812
QY 1294 GGCCCAATGTCTTCTGTATGTCTGACACATGCTGTGGGCTTGTGGGCTTGTGGGCTG 1353
Db 1813 GGCCCAATGTCTTCTGTATGTCTGACACATGCTGTGGGCTTGTGGGCTTGTGGGCTG 1872
QY 1354 TACGTGGGGCGCCCAAGTCTCCACCTGCTTGGCCGACGCTTGTCTTCCCTTTC 1413
Db 1873 TATGTGGGGCGCCCAAGTCTCCACCTGCTTGGCCGACGCTTGTCTTCCCTTTC 1932
QY 1414 TTCCATTTGCAATCTCTGTATGTCTGAGGCTTGTCTTCCAGATGCTGTGGCTTCAAG 1473
Db 1933 TTCTCCATTTGCAATCTCTGTATGTCTGAGGCTTGTCTTCCAGATGCTGTGGCTTCAAG 1992
QY 1474 ATGCCAGGCGCTTCCAGGCGCTTACAGTACTGGGCTGCTTCCAGGCGCTTCCAGT 1533
Db 1993 ATGCCAGGCGCTTCCAGGCGCTTACAGTACTGGGCTGCTTCCAGGCGCTTCCAGT 2052
QY 1534 TCTATGCAATTTATCAGGTAATCAAAATGCTATTTGTGATTTGGCATGCTGGCAGG 1593
Db 2053 TTTCTGTCCTTTCATCAGGCGCTTCAAGGCTGCTTGGGCTGCTGGCAGCATGCTGGCAGC 2112
QY 1594 CCTAGTCCCAAGC---CCGTAAGTACCCGCTTACAGTACTGGGCTGCTTCCAGT 1650
Db 2113 ACATCAACACCCATTTGGCGGAGACCGGAGTACCCCAATCATGATCTCTCTGCTG 2172
QY 1651 AACCCCACTACCCCAAGAGCTGCTTCAACACAGCTGAGCTGCTGCTCTCAAGT 1710
Db 2173 CACCTTACTACCCCAAGAGCTGCTTCAACACAGCTGAGCTGCTGCTCTCAAGT 2232
QY 1711 GTGGGTTTCAAGCTTCCATGAGGCAAGAGCTGCTGCTCAACATCAACAGAGCCAG 1770
Db 2233 CTGGGTTTCAAGCTTCCATGAGGCAAGAGCTGCTGCTCAACATCAACAGAGCCAG 2292
QY 1771 TTCAATCAACCTTACATGAGCTTCTATTTCACTCATCCGCTGCTTCTGACCTTCA 1830
Db 2293 TTCAATCAACCTTACATGAGCTTCTTCTTCACTCATCCGCTGCTTCTGACCTTCA 2352
QY 1831 TCTGCTCAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1890
Db 2353 TCTGCTCAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2412
QY 1891 CTGGCCATCAAGCTTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1950
Db 2413 CTGGCCATCAAGCTTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2472
QY 1951 CGCAACAGCGCGGCTTCTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2004
Db 2473 CGCAACAGCGCGGCTTCTCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2526

```

RESULT 4
 AF337041 3060 bp mRNA linear ROD 25-Apr-2001
 LOCUS AF337041
 DEFINITION Mus musculus candidate taste receptor T1R2 (T1R2) mRNA, complete cds.

```

ACCESSION AF337041
VERSION AF337041.1 GI:13785658
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Montmayeur, D.P., Liberles, S.D., Matsumami, H. and Buck, L.B.
TITLE A candidate taste receptor gene near a sweet taste locus
JOURNAL Nat. Neurosci. 4 (5), 492-498 (2001)
MEDLINE 21219400
PUBMED 11319557
REFERENCE
2 (bases 1 to 3060)
AUTHORS Montmayeur, D.P., Liberles, S.D., Matsumami, H. and Buck, L.B.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-2001) Neurobiology, Harvard Medical School, 220 Longwood Avenue, Boston, MA 02115, USA
FEATURES
SOURCE
1..3060
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="4"
1..3060
/gene="T1R2"
32..2563
/note="T1R2"
/note="G protein coupled receptor"
/codon_start=1
/product="candidate taste receptor T1R2"
/db_xref="GI:13785659"
/protein_id="AAK39438.1"
Query Match 60.7%; Score 1220.6; DB 10; Length 3060;
Best Local Similarity 76.5%; Pred. No. 9.8e-192;
Matches 1560; Conservative 0; Mismatches 449; Indels 30; Gaps 4;
QY 2 TCACCTACAGGCGCATCAGCATGAGCTGCGAGACAAAGTGGCTTCCGCGCTTGTGCTG 61
Db 525 TCACATATAGCGCCATCACCAAGCTGCGAGACAAAGCGGCTTCCCTGCTGCTGCTG 584
QY 62 GTACACACCCAGGCGGCGAGCAGCAGTGCAGGCGCATGCTGAGTGTGCTGCACTTCC 121
Db 585 GCACTGTGCCAGGCGGCGAGCAGCAGTGCAGGCGCATGCTGCACTGATGCTGCTG 644
QY 122 GCTGGAATGATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 181
Db 645 AGTGAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 704
QY 182 TGCTTGAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 238
Db 705 TGCTTGAGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 764
QY 239 CCACATGCGACCCCAACAGCAATGAGCTGAGAGAGCGGCGCTGCTGCAATTTG 298

```

D	b	765	CTGTACCCAGAACCCACACAGGCCGTGTGAGGCTGTGAGGAGCAGGACCAACTGTGACAACTATCC	824
Q	y	299	TGGACAACTCTGACAGACAGCAACGCGCGCTGTGCTGTCTCTGCTCTGCCCCAGCTGACCC	358
D	b	825	TGGAAACACTCTCGCGGAGACCTCGCGCGTGTGTGTGTATATCTCTGCCAGAGCTGAGCC	884
Q	y	359	TGTACCACTTTCTTAATGAGGTGTGTGGCAGAACTTCAACGGGCGCGGTGTGTGTGTGCT	418
D	b	885	TGCACAACTTTCTTCCGGGAGGTGTGTGGCTGTGAACTTCAACAGGCTTTGTGTGTGTGCT	944
Q	y	419	CCGAGTCTCTGGGCCCATGACCCGGTCTCTGTGCACAACCTCAACGAGAGCTGTGGCCACTTGGACA	478
D	b	945	CTGAGTCTCTGGGCCCATGACCCCTTTCTTACAACTCAACAGAGCTGTGGCCACACAGGGCA	1004
Q	y	479	CCTTCTCTGGGCATCAACATCCAGAGCGTCCCATCCCGGGCTTCAGTAGTATCCGCGAGT	538
D	b	1005	CTTTCTCTGGGCTCAACATCCAGAGGCTGTGCATCTCTGGCTTCAGCCAGTTTCCAGTGC	1064
Q	y	539	GGGGGCCCAAGGCTGTGGGCGCGCAACCCCTCAAGCAGAACACAGCCAGAGCTTACTCTGCAAC	598
D	b	1065	GCCACGACAAACCCAGAGATCCATGCTATAGACAGACACACCCCTGAGAGCTACTCTTAAC	1124
Q	y	599	AGGATGTGGACAAATCGCTGAAACCCCACTTGTCTTCAACACACATCTCAGAGCTGTG	658
D	b	1125	AGGATGTGAGCGCTCGCATGAACTCAACCAAGTCTCTTTAACAAAGCTTCTCACTGCTTTGG	1188
Q	y	659	GGGAGCGTGTCTGTACAGCGTGTACTCTGCGGCTATGCTGTGGCCATGCCCTGCACA	718
D	b	1185	GGGAGCGTGTGTGTACAGTGTGTACTCGGCGCTGTACGGGAGTGAAGCCACACCTTCCACA	1244
Q	y	719	GGCTCTCTGGGTGTGACAAAAGCAACCTGCACCAAGAGGCTGTACCCCTGGCAGCTGC	778
D	b	1245	GACTCTCTCACTGCACACAGGCTCGCGTGCACCAAGCAAACTGCTATCCATGGCAGCTAC	1304
Q	y	779	TTGAGGAGATCTGGAAGTCAACTTCACTCTCTGTGACCAACAATCTTCTTGCAGCCG	838
D	b	1305	TCAGGAGATCTGGCAGTAGTCAACTTCAAGCTCTCTGGGCAACAAGCTCTTCTTGCAGAAC	1364
Q	y	839	AAGGGAGAGTGTGGCTCTGCACTTGGAGATTTCAGTGTGCATGTGGACCGGAGCCAGAAATC	898
D	b	1365	AAGGGAGATCCCATCTCTCTGTGCATCTCAAGTGGCAATGGGGCTGTGAGCCAGAAAC	1424
Q	y	899	CCTTCCAGAGGTGGCGCTCCACTACCCCTGAGGAGCAAGCTGGAAGAAACAT---CAAGA	955
D	b	1425	CCTTCCAAAGCATGTGCTCTACTCTCCCAACGAGACAGAGGCTGACCTTCAATTAGCAATG	1484
Q	y	956	CATCTGTGCACACCGTCGACAAACACAGATCCCTATGTCTCATGTGTTCAGAGAGTCCCACT	1015
D	b	1485	TGTCTGTGTACACCCCAACACAGAGTCCCATATCTATGTCTTAAAGATTGTCCAGC	1544
Q	y	1016	CAGGGCAAAAGAGAGAGCGTGTGGGCAATCAAGTGTGTCTTCAGTGCATGACGTCC	1075
D	b	1545	CTGGGCAAAATGAAAAAACCATATAGGCCCTCCACCCGTGTCTGAGTGTGGAGCTGTC	1604
Q	y	1076	TTCCCGGACCTTCTCAACCAACATGAA-----TGCCCGAATA	1114
D	b	1605	CGCGGGGACCTTACTCAACCGATCAAGTAAATGATTAACTGTCTGTCTCTCCCGGGTT	1664
Q	y	1115	ACGATGTCTCTTACAGAGTGAAGCCCTCGCTTCAAGGGGAGCGGTCTTCGAGAT	1174
D	b	1665	CCATGTGTCTTACAAACAAACATGTGCTTCTTCAAGGGGGGCTGTCTTCGTGGAGT	1724
Q	y	1175	GGCATGAGGACCCACCATGCTGTGTGGCCCTGTGTGGCCGCTCGGGGCTCTCAACACC	1234
D	b	1725	GGCAGCAAGTCCCATATGTGTGTGACATCTGTGGCGCCCTCGGGCTTCATCAATAGC	1784
Q	y	1235	TGGCATCTCTGTGTATATTCTGAGGCACTTCCAGACCCCATATAGTGTGCTGGCTGGGG	1294
D	b	1785	TGGCATCTCTGTCTCATCTTCTGGGAGCAATTTCCAGACGCCCATGTGTGCTGTGGCGCG	1844
Q	y	1295	GGCCCATGTGCTTCTCATGCTCAACACTGTGTGTGTGTCATATAGTGTGTGCGGCTGT	1354
D	b	1845	GGCCCATGTGCTTCTCATGCTGTGTGCGCTGTGTGTGCTGTGCGGATGTGTGCTTCCCGCTGT	1904

[illegible]

[illegible]

REFERENCE 1 (bases 1 to 2532)
 AUTHORS Nelson, G., Hoon, M.A., Chandrasekar, J., Zhang, Y., Ryba, N.J. and Zaker, C.S.
 TITLE Mammalian sweet taste receptors
 JOURNAL Cell 106 (3), 381-390 (2001)
 MEDLINE 21400448
 PUBMED 11509186
 REFERENCE 2 (bases 1 to 2532)
 AUTHORS Hoon, M.A., Mueller, K.L., Zhang, Y., Adler, E., Ryba, N.J.P. and Zaker, C.S.
 TITLE Direct Submission
 JOURNAL Submitted (15-APR-2001) Taste and Smell Unit, NIDCR, NIH, 10 Center Drive, Bethesda, MD 20892-1188, USA
 FEATURES
 source 1. 2532
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="4"
 /map="69-71 cm"
 1. 2532
 /note="G-protein coupled receptor (GPCR)"
 /product="sweet taste receptor T1R2"
 /protein_id="AAK51604.1"
 /db_xref="GI:15147681"
 /translation="MGQPARTLHLFLHLALPKPVMVLVNSDFHLADYLLGLFTLHANVSVSHSLQVPCNKNEYMKVLYMLQAMRAVEINNCSSILGLGYEMVDYLSNNTIOPGLYELSDIDFLPIKDYRPOVAVYIGDNSLISLSELYPOVYSAITPKIDKREPRAMLRTPSRATHHEMOVALHFGNNATLVISDDYGRENSHLQRLNTGSDICIAQOVLVPEPNOAYRPEODLDLRLKRSARVVIPEPELSLHNEFREVLNRTGFWIAISEWADIVLNLRLRTGFLVTLQVSIPEQSQFVRHDKRPMNETSLRTTCNODCAIDMNTESNNVLMISGEHVYVSYAVAYAVATLRLHLCNQVCKQIVYPMOLREIWHNFTLGNLPFDEGDPMIADIOMWGLSONFPOSJASYSPETRLTYISNVMYPPNMTVPISMCSQPGOMKKPIGHPCCEFCVDCPDPTLYNRSVDEENQSCPSGMSRYNNIACFEKRLAFLEMH EPTVYVITLALGRISTALIIILIRHRIQTEVRSAGPKCFILVPIILAFAMPVYVGPVESCFCRQAFYVCFVSCISCTVRSQIVCVKMARILPSAIGFPMRTHGPYVFATFTVAALVAGNMILATTINIPIDPDPIIILSCHPNRDLNFTSMDL LLSVAFSAVYVKGELPTYNENAKFTLSDMTSFTSSILCTEMSHVDCILVYIMDL VYVNLFIALIGLGFPGKCYMILFPERPNTSAVENSMIGQYMKRS"

BASE COUNT 498 a 824 c 631 g 579 t
 ORIGIN

Query Match 60.3%; Score 1212.6; DB 10; Length 2532;
 Best Local Similarity 76.3%; Pred. No. 2,1e-190;
 Matches 1555; Conservative 0; Mismatches 454; Indels 30; Gaps 4;

QY 2 TCACCTACAGCGCCATTCAGCATGAGTGGGAGACAAAGTGGCGCTTCCGGCTTTGCTGC 61
 DB 494 TCACATATACGCGCATACCGGACAAAGTGGCAAGAGCGGCTTCCCATGCTGC 553
 QY 62 GTACACACCCAGCGCGGACGACGATCGAGGCGCATGTGACAGTCTGCACTTCC 121
 DB 554 GCACGTGCCGACGCGGACCCACCAATGAGGCGCATGTGCAACTGATGTTCACTTCC 613
 QY 122 GCTGGAGTCTGATCTGTCTGTGTAGAGCGACACCTTATGGCGCGCAATATGGCCAGC 181
 DB 614 AGTGGAACTGTGATCTGTGTCTGTGTAGCATATTTATGGCGCGAAGAACGACCC 673
 QY 182 TGGTTGGCGAGCGGTGGC---CGGCGCGACATCTGATCGGCTTCAGAGAGCGGTGC 238
 DB 674 TGTGTGACGACGCTCTGACCAACATCTGCAATCTGCTTCCAGAGAGTTTCTGC 733
 QY 239 CCACACTGACGCCAACACAGACATGACGTACAGAGAGCGCCAGCGCTTGTACCAATTG 298
 DB 734 CCGTACAGAAACCAACAGAGGTGTAGGCTGAGGAGGACAGCAACTGGACAACTCC 793
 QY 299 TGGACAGAGTGTGACAGACAGACGCGGCTGTGTCTGTCTGTGCGCGACCTGACCC 358
 DB 794 TGGACAGAGTGTGCGGCGGATCTGTGTGTGTGTGTATTTCTCCGAGACTGAGCC 853
 QY 359 TGTACACTTCTTCATGAGT 418

DB 854 TGCACAACCTTTCCTCGTAGAGT 913
 QY 419 CCGAGTCTCGGGCCATCGACGACCGGCTCTGTGCAACCTTCACGAGAGTGGCCACTTGGGCA 478
 DB 914 CTGAGTCTCGGGCCATCGACGACCGGCTCTGTGCAACCTTCACGAGAGTGGCCACTTGGGCA 973
 QY 479 CTTTCTCGGGCCATCGACGACCGGCTCTGTGCAACCTTCACGAGAGTGGCCACTTGGGCA 538
 DB 974 CTTTCTCGGGCCATCGACGACCGGCTCTGTGCAACCTTCACGAGAGTGGCCACTTGGGCA 1033
 QY 539 GGGGCCCAAGGCTGGGCGGCCACCCCTGACAGAGACCGACAGACTATACCTGCAACC 598
 DB 1034 GCCATGCAAGCCAGGATATGCGATGCTTAACGACACCGACCTGCGGACTCTGTAAAC 1093
 QY 599 AGGAGTGGCAACATGCGCTTACGACCGACCGACCTGCTTCAACACCATTTCTAGGCTGTG 658
 DB 1094 AGGAGTGGCAACATGCGCTTACGACCGACCGACCGACCTGCTTCAACACCATTTCTAGGCTGTG 1153
 QY 659 GGGAGCGTGTCTGTACAGCGTACTCTGCGGTCTATGCTGTGGCCATGCGCTGACCA 718
 DB 1154 GGGAGCGTGTCTGTACAGCGTACTCTGCGGTCTATGCTGTGGCCATGCGCTGACCA 1213
 QY 719 GCTTCTCGGCTGTGACAAAGACCTGACCAAGAGAGGTGTCTACCTGGCAGCTGC 778
 DB 1214 GACTCTTCCATCTCAATCAAGTCCGCTGACCAAGCAAAATGCTATCAATGACGCTAC 1273
 QY 779 TTGAGGAGATCTGGAAGTCAATCTCACTCTCTGACCAACCAATCTTCTGACCCGC 838
 DB 1274 TCAGGAGATCTGGAAGTCAATCTCACTCTCTGACCAACCAATCTTCTGACCCGC 1333
 QY 839 AAGGGAGCTGTCTGTGCACTTGGAGATTGTCCAGTGTGCAATGGAGCGACCGACAGANTC 898
 DB 1334 AAGGGAGCTGTCTGTGCACTTGGAGATTGTCCAGTGTGCAATGGAGCGACCGACAGANTC 1393
 QY 899 CTTTCAGAGCGTGGCTCTCTACTACCCCTGAGGAGACGTAAGAATAT---CAGA 955
 DB 1394 CTTTCAGAGCGTGGCTCTCTACTACCCCTGAGGAGACGTAAGAATAT---CAGA 1453
 QY 956 CATCTGTGACACCGTTCACACAGATCCCTTATGTCATGTGTTCACAGAGGTGCGAGT 1015
 DB 1454 TGTCTGTGATACCCCAACACAGAGTCCCATATGCTGTTCACAGAGGTGCGAGT 1513
 QY 1016 CAGGGCAAAAGAAAGACCTGTGGGCAATCCAGTCTGTCTGTGAGTGTGAGTGTGCGC 1075
 DB 1514 CTTGGGCAAAAGAAAGACCTGTGGGCAATCCAGTCTGTCTGTGAGTGTGAGTGTGCGC 1573
 QY 1076 TTTCCGCGACCTCTCTCAACACACATGA-----TGGCCGATA 1114
 DB 1574 CCGCGGACACCTCTCTCAACACACATGA-----TGGCCGATA 1633
 QY 1115 ACGAGTGTCTTACAGAGTGAACCTCTGTCAAGCGGACGTGTCTCTGTGAAT 1174
 DB 1634 CCAATGTGTCTTACAGAGTGAACCTCTGTCAAGCGGACGTGTCTCTGTGAAT 1693
 QY 1175 GGCATGAGGCAACCAACATGCTGTGGCGCTGTGGCGCGGCTGTCTCTCAAGACCC 1234
 DB 1694 GGCATGAGGCAACCAACATGCTGTGGCGCTGTGGCGCGGCTGTCTCTCAAGACCC 1294
 QY 1235 TGGCCATCTCTGTGTATTTCTGAGAGCACTTCCAGACACCAATAGTTCTCGGCTGGG 1294
 DB 1754 TGGCCATCTCTGTGTATTTCTGAGAGCACTTCCAGACACCAATAGTTCTCGGCTGGG 1813
 QY 1295 GCCCATGTCTCTGTGTATTTCTGAGAGCACTTCCAGACACCAATAGTTCTCGGCTGGG 1354
 DB 1814 GCCCATGTCTCTGTGTATTTCTGAGAGCACTTCCAGACACCAATAGTTCTCGGCTGGG 1873
 QY 1355 ACGTGGGCGCGCCAGAGTTCACACCTGCTGTGGCGCGGCGGCTTCTCCCTGCTGCT 1414
 DB 1874 AATGTGGCGCGCGCCAGAGTTCACACCTGCTGTGGCGCGGCGGCTTCTCCCTGCTGCT 1933
 QY 1415 TCACATTTGATCTCTGTATGCGCGTGTGCTTTCAGATGATGCTGTGCGCTTCAAGA 1474
 DB 1934 TGTCTGTCTCTCTGTATGCGCGTGTGCTTTCAGATGATGCTGTGCGCTTCAAGA 1993

QY 1475 TGGCAGCGGCTGCCAGCGGCTACAGCTACTGGGCTCGGCTACCGAGGGCCCTACGCT 1534
 Db 1994 TGGCCAGAGCGCTGCGACGAGCGGCTTCTGTGATCGTTACACGCGGCTACGCT 2053
 QY 1535 CTATGCGATTATACAGGCTACTCAAAATGTCATGTGTATGTCATGTCGACGCG 1594
 Db 2054 TCGTGGCCTTATCATGAGCGCGCTGAGGCTGCTGGCGGCGACATGCTGGCCACCA 2113
 QY 1595 CTGATGCCACCC---CCGTAAGACCCCGATGACCCCAAGATGACATGTCCTCTGTA 1651
 Db 2114 CCATCAACCCCATGGCGCGGACCGGACCGGATGACCCCAATATCATCTCTCTGCC 2173
 QY 1652 ACCCAATACCGCAACAGCTGTGTCATCAACACAGCGGCTGCTGCTCATGTTG 1711
 Db 2174 ACCCAATACCGCAACAGCGGCTGCTGTCATCAACACAGCTGCTGCTGCTGCTG 2233
 QY 1712 TGGGTTTCAAGCTGCGCTACATGAGGCAAGAGCTGCGCCACCACTACACGAGCCAA 1771
 Db 2234 TGGGTTTCAAGCTGCGCTACATGAGGCAAGAGCTGCGCCACCACTACACGAGCCAA 2293
 QY 1772 TCATCACCTTACAGCTGCTTCTTATTTTCACTTCACTGCTGCTGCTGCTGCTGCT 1831
 Db 2294 TCATCACCTTACAGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2353
 QY 1832 CTGCTACAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1891
 Db 2354 CTGCTACAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2413
 QY 1892 TGGCATCAGGCTGCGCTTCTGCGCCCAAGTGTCTACATGCTCTTCTTCTTCTTCT 1951
 Db 2414 TGGCATCAGGCTGCGCTTCTGCGCCCAAGTGTCTACATGCTCTTCTTCTTCTTCT 2473
 QY 1952 GCACACGCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2010
 Db 2474 GCACACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2532

RESULT 7

AF458149S6

LOCUS AF458149S6 930 bp DNA linear PRI 12-APR-2002

DEFINITION Homo sapiens taste-specific G-protein coupled receptor T1R2

ACCESSION AF458154

VERSION AF458154.1 GI:20142332

KEYWORDS

SEGMENT

SOURCE

ORGANISM

Homo sapiens.
Homo sapiens

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SOURCE

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

gene

AF458151.1:1..774,AF458152.1:1..210,AF458153.1:1..123,
 1..>930)
 /gene="TAS1R2"
 /product="taste-specific G-protein coupled receptor T1R2"
 join(AF458149.1:2..183,AF458150.1:3..303,
 AF458151.1:1..774,AF458152.1:1..210,AF458153.1:1..123,
 1..>930)
 /gene="TAS1R2"
 /note="T1R2: putative human sweet taste receptor"
 /codon_start=1
 /product="taste-specific G-protein coupled receptor T1R2"
 /protein_id="AA012233.1"
 /db_xref="GI:20142333"
 /translation="MGPRKATIGSLFLLWLAIPAENSDFYLPDYLGLGSLFSLHAN
 MKGIVHNLQVLPKREYEVYICYNLMQANRPFVERLNDSSILPVLGLYELVDC
 YISNVPVLFTLHEDNLEIQEDISYISRVAVIGPDNSESVMVAVNLSFLPLP
 QITYSAISDELARDVREPALLRTTPSADHNEAVOQLMEHVMVILVLSDDYGRD
 NGOLGERSVARDICIAFOETPLTQDNHSEERLTVYDKLQDSFARVYVFS
 PDLTYHFNENLVRFNFTGAVWIASMAIDPVNLNTELGHTFLGITIOSYPIG
 FSEFRMGPOGPPPLSRTSOSYCNOCNCLNATLSPNTLRLSGRYYSVYSAV
 YAVAHALSLGCKSTCKRYVYPMOLBEIKVNFLLDHOIFPDQGVALHLET
 VOWMDRSNPFQSVASTYPLQRLKNIQDISHTVNTTIPMSKSRKSGOKKKPV
 GIHYCEECIDLPGLNLTEDYECQACPNEMSYQSETSCERDLVLEMEHAP
 IAVALLAELGLSTLALIVIFMRHPTPIYASAGPMCFMLTLVAVAVYVYGP
 PKVSTLCROALFPLCFCTICISIAVSFOVCAFKMSRPRAYSVWRQGYVM
 AFITVLKMTIVYIGMLATGLSPTRPDDEKITIVGCPNRYRSLFENSLDLTSLV
 VGSFPAIMGKELPNTYNEAKITISMTFIFSSVSLCTPMAYSGVLTVIYDLVTVL
 NLATISLGYEPKCYMILFPERNTPAIFNSMIGYTRRD"
 1..>930
 /gene="TAS1R2"
 /number=6
 BASE COUNT 171 a 322 c 222 g 215 t
 ORIGIN

Query Match 43.8%; Score 879.8; DB 9; Length 930;

Best Local Similarity 99.4%; Pred. No. 1.8e-133;

Matches 904; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1105 TGGCCGAATACGAGTGGCTCCACAGAGTAGACTCTCTGTTCAAGCGGAGCTGTC 1164
 Db 22 TGGCCGAATACGAGTGGCTCCACAGAGTAGACTCTCTGTTCAAGCGGAGCTGTC 81
 QY 1165 TTCCGTGAATGAGCAGAGGACACCAACATGCTGAGCCCTGAGCGGCTGAGCTTC 1224
 Db 82 TTCCGTGAATGAGCAGAGGACACCAACATGCTGAGCCCTGAGCGGCTGAGCTTC 141
 QY 1225 CTCAGCACCTGGCCATCTGCTGATATTCTGAGGACATTCAGACACCACTAAGTGGC 1284
 Db 142 CTCAGCACCTGGCCATCTGCTGATATTCTGAGGACATTCAGACACCACTAAGTGGC 201
 QY 1285 TGGGCTGGGGGCCCCATGCTCTCTGATGCTGACACGCTGCTGCTGCTGCTGCTGCT 1344
 Db 202 TGGGCTGGGGGCCCCATGCTCTCTGATGCTGACACGCTGCTGCTGCTGCTGCTGCT 261
 QY 1345 GTCCCGGTGATGCTGGGGGCGGCAAGGCTCTCACCTGCTGCTGCTGCTGCTGCTT 1404
 Db 262 GTCCCGGTGATGCTGGGGGCGGCAAGGCTCTCACCTGCTGCTGCTGCTGCTGCTT 321
 QY 1405 CCCCTGCTGCTTACAAATTTGATCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1464
 Db 322 CCCCTGCTGCTTACAAATTTGATCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 381
 QY 1465 GCGTCAAGATGCGCAGCGGCTTCCACAGCGGCTACAGTATGCTGCTGCTGCTGCTGCT 1524
 Db 382 GCGTCAAGATGCGCAGCGGCTTCCACAGCGGCTACAGTATGCTGCTGCTGCTGCTGCT 441
 QY 1525 CCTTACGCTCTATGAGCTTTATACAGGCTACCAAAATGCTATGCTGCTGCTGCTGCT 1584
 Db 442 CCTTACGCTCTATGAGCTTTATACAGGCTACCAAAATGCTATGCTGCTGCTGCTGCT 501
 QY 1585 CTGGC--ACGGCTTACGCTCCAC--CCGTAAGACCCCGATGACCCCAAGATCAAT 1641
 Db 502 CTGGCAGCGGCGCTACGCTCCACCGCTACTGACCCCGATGACCCCAAGATCAAT 561

QY 1942 TACCGGAGCGCAACGCCGCCCTACTTCAACAGCATGATCCAGGCTACACCATGAG 2001
 |||||||
 Db 192693 TACCGGAGCGCAACGCCGCCCTACTTCAACAGCATGATCCAGGCTACACCATGAG 192634
 |||||||
 QY 2002 AGGAGCTAG 2010
 |||||||
 Db 192633 AGGAGCTAG 192625
 |||||||
 RESULT 9
 AF458149S3
 LOCUS Homo sapiens taste-specific G-protein coupled receptor 11R2
 DEFINITION Homo sapiens taste-specific G-protein coupled receptor 11R2
 (TAS1R2) gene, exon 3.
 ACCESSION AF458151
 VERSION AF458151.1 GI:20142329
 KEYWORDS
 SEGMENT
 SOURCE
 ORGANISM
 Homo sapiens.
 3 of 6
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 774)
 Li,X., Staszewski,L., Xu,H., Durick,K., Zoller,M. and Adler,E.
 Human receptors for sweet and umami taste
 JOURNAL
 Proc. Natl. Acad. Sci. U.S.A. 99 (7), 4692-4696 (2002)
 MEDLINE
 21927605
 PUBMED
 11917125
 REFERENCE
 2 (bases 1 to 774)
 Li,X., Staszewski,L. and Adler,E.
 Direct Submission
 JOURNAL
 Submitted (14-DEC-2001) Genomix, Inc., 11099 N. Torrey Pines Road,
 La Jolla, CA 92037-1007, USA
 FEATURES
 source
 1..774
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /gene="TAS1R2"
 /number="3"
 exon
 1..774
 BASE COUNT 145 a 268 c 222 g 139 t
 ORIGIN
 Query Match 38.5%; Score 774; DB 9; Length 774;
 Best Local Similarity 100.0%; Pred. No. 5.1e-118;
 Matches 774; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 361 TACCACTTCTCAATAGAGTGTGCGCCAGACTTCACGCGCGCGTGTGATGCGCTCC 420
 |||||||
 Db 361 TACCACTTCTCAATAGAGTGTGCGCCAGACTTCACGCGCGCGTGTGATGCGCTCC 420
 |||||||
 QY 421 GAGTCCTGGGCGCATGACCGCGCTCTGACACATCTACAGAGACTGAGGCATCTGGGACAC 480
 |||||||
 Db 421 GAGTCCTGGGCGCATGACCGCGCTCTGACACATCTACAGAGACTGAGGCATCTGGGACAC 480
 |||||||
 QY 481 TTCCTGGGCGCATGACCGCGCTCTGACACATCTACAGAGACTGAGGCATCTGGGACAC 540
 |||||||
 Db 481 TTCCTGGGCGCATGACCGCGCTCTGACACATCTACAGAGACTGAGGCATCTGGGACAC 540
 |||||||
 QY 541 GCGCCACAGAGCTGGGCGCGCCACCTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 |||||||
 Db 541 GCGCCACAGAGCTGGGCGCGCCACCTCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
 |||||||
 QY 601 GAGTCGACAACTGCTGACAG 660
 |||||||
 Db 601 GAGTCGACAACTGCTGACAG 660
 |||||||
 QY 661 GAGTCGCTGCTGTACAG 720
 |||||||
 Db 661 GAGTCGCTGCTGTACAG 720
 |||||||
 QY 721 CTCCTCGGCTGTGACAAAGACACCTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 774
 |||||||
 Db 721 CTCCTCGGCTGTGACAAAGACACCTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 774
 |||||||
 RESULT 10
 AL831790
 LOCUS Mus musculus chromosome 4 clone RP23-334I5, *** SEQUENCING IN
 DEFINITION
 ACCESSION AL831790.2 GI:22416181
 VERSION
 KEYWORDS
 HTG; HTGS_PHASE2; HTGS_ACTIVEPIN; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE
 house mouse.
 ORGANISM
 Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 1 (bases 1 to 206074)
 Wall,M.
 JOURNAL
 Direct Submission
 Submitted (16-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Aug 21, 2002 this sequence version replaced gi:21727737.
 COMMENT
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquerry@sanger.ac.uk
 Contact: humquerry@sanger.ac.uk
 Project Information
 Center project name: bm33415
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Chemistry: dye-terminator; 100% of reads
 Consensus quality: 205915 bases at least Q40
 Consensus quality: 206010 bases at least Q30
 Consensus quality: 206049 bases at least Q20
 Insert size: 206074; sum-of-contigs
 Insert size: 220515; 2.7% error; agarose-fp
 Quality coverage: 7.94x in Q20 bases; sum-of-contigs quality
 coverage: 7.73x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 Location/Qualifiers
 1..206074
 /organism="Mus musculus"

```

/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-33415"
/clone_1ib="RP23-33415"
misc_feature 1..206074
            /note="assembly_fragment:00142"
BASE COUNT 48675 a 52290 c 53898 g 51211 t
ORIGIN
Query Match      30.1%; Score 605; DB 2; Length 206074;
Best local Similarity 79.9%; Pred. No. 2e-90;
Matches 726; Conservative 0; Mismatches 180; Indels 3; Gaps 1;

QY 1105 TGCCTGGAATACGAGTGTCTTACAGAGTGAAGCTCTCTTCAAGCGGACCTGTC 1164
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52178 TGCCCGGGTTCATGTGTCTTACAGAGTGAAGCTCTCTTCAAGCGGACCTGTC 52237
QY 1165 TTCTGTGAATGAGTGAAGGACCCACCATGCTGTGCGCTGCGCGCGCTGTC 1224
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52238 TTCTGTGAATGAGTGAAGGACCCACCATGCTGTGCGCTGCGCGCGCTGTC 52297
QY 1225 CTCAGCACCTGAGGACCTGTGTATCTGTGAGGACCTTCCAGACCATGTCGC 1284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52298 ATCAGTACGCTGTGAGTGTCTGTCTGTGAGACATTTCCAGACCATGTCGC 52357
QY 1285 TCGGCTGGGGCCCATGTCTTCTGTGATCTGTGACACTGCTGTGAGCATGTCG 1344
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52358 TCGGCTGGGGCCCATGTCTTCTGTGATCTGTGACACTGCTGTGAGCATGTCG 52417
QY 1345 GTCCCGGTGTACGTGGGGCCCGCCAGGCTCCACCTGCTGCGCGCGCTTTC 1404
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52418 GTCCCGGTGTACGTGGGGCCCGCCAGGCTCCACCTGCTGCTGCGCGCTTTC 52477
QY 1405 CCCCTGTCTTCAATTTGATGATCTGTATGCGCGCTGCTTCCAGATGTCGTC 1464
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52478 ACCGTTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 52537
QY 1465 GCTTCAAGATGGGACCGCTTCCAGGCGCTTCCAGTACTGCTGCTGCTGCTGCT 1524
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52538 GCTTCAAGATGGGACCGCTTCCAGGCGCTTCCAGTACTGCTGCTGCTGCTGCT 52597
QY 1525 CCCCTGTCTTCAATTTGATGATCTGTATGCGCGCTGCTTCCAGATGTCGTC 1584
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52598 CCCCTGTCTTCAATTTGATGATCTGTATGCGCGCTGCTTCCAGATGTCGTC 52657
QY 1585 CTGAGCAGGCGC---TCAGTCCACACCCCTGATGACCCCGATGACCCCAAGATC 1641
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52658 CTGAGCAGGCGC---TCAGTCCACACCCCTGATGACCCCGATGACCCCAAGATC 52717
QY 1642 GTCTCTGTAAACCCCACTACCGGACGCTGCTGTTCACACACGCTGAGCTGTC 1701
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52718 GTCTCTGTAAACCCCACTACCGGACGCTGCTGTTCACACACGCTGAGCTGTC 52777
QY 1702 CTCTCACTGTGTGCTTACCTTCTGCTTACATGAGGCAAGAGCTGCGCCCACTA 1761
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52778 CTCTCACTGTGTGCTTACCTTCTGCTTACATGAGGCAAGAGCTGCGCCCACTA 52837
QY 1762 GAGGCAAGTTCATCACCCCTGAGCATGACCTTCTATTTACCTGCTGCTGCTG 1821
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52838 GAGGCAAGTTCATCACCCCTGAGCATGACCTTCTATTTACCTGCTGCTGCTG 52897
QY 1822 ACCCTTCACTGTGTGCTTACCTTCTGCTTACATGAGGCAAGAGCTGCGCCCACTA 1881
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52898 ACCCTTCACTGTGTGCTTACCTTCTGCTTACATGAGGCAAGAGCTGCGCCCACTA 52957
QY 1882 CTCAACCTCTGTGCGCATGACCTGCTGCTTACATGAGGCAAGAGCTGCGCCCACTA 1941
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52958 CTCAACCTCTGTGCGCATGACCTGCTGCTTACATGAGGCAAGAGCTGCGCCCACTA 53017
QY 1942 TACCCGAGGCAACAGCGCGCTTACTTCAAGCATGATGATGAGGCTACACATGAGG 2001
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 53018 TACCCGAGGCAACAGCGCGCTTACTTCAAGCATGATGATGAGGCTACACATGAGG 53077

```

QY 2002 AGGAGCTAG 2010
 Db 53078 AAGAGCTAG 53086

RESULT 11
 AC115397/c
 LOCUS
 DEFINITION
 *** 69 unordered pieces.
 AC115397
 AC115397.2 GI:21737066
 HTG: HTGS_PHASE1.
 Norway rat.
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE AUTHORS

1 (bases 1 to 154891)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alshrocks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barbara,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
 Bunay,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,
 Carion,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
 Homs,E., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtarge,O., Liu,C., Liu,J., Liu,W., Louisedge,H.,
 Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhney,E., McLeod,M.P., Meadow,M., Mel,G., Metzger,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabat,K., Morgan,M., Morris,S.,
 Moser,N., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokoko,S., Oguh,M., Okunoye,G.,
 Oragunye,N., Oyiedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,K., Ren,Y.,
 Rives,M., Rojas,A., Rojudoan,I., Rolfe,M., Ruiz,S., Saverly,G.,
 Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
 Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabori,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
 Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Nelson,D.,
 Weinstein,G. and Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 154891)
 Worley,K.C.
 Direct Submission
 Submitted (19-MAR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 154891)
 Worley,K.C.
 Direct Submission
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Jul 12, 2002 this sequence version replaced gi.19549101.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: CNRS

Center clone name: CH230-253G20

----- Summary Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 77784 bases at least Q40

Consensus quality: 77784 bases at least Q30

Consensus quality: 81166 bases at least Q20

----- NOTE: Estimated insert size may differ from sequence length

(see <http://www.hgsc.bcm.tmc.edu/docs/genbank.draft.data.html>).

NOTE: This is a 'working draft' sequence. It currently

consists of 69 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 1066: contig of 1066 bp in length
1067 1166: gap of unknown length
1167 2409: contig of 1243 bp in length
2410 2509: gap of unknown length
2510 3864: contig of 1355 bp in length
3865 3964: gap of unknown length
3965 5080: contig of 1116 bp in length
5081 5180: gap of unknown length
5181 6333: contig of 1153 bp in length
6334 7893: gap of unknown length
7894 7994: gap of unknown length
7994 9421: contig of 1428 bp in length
9422 9521: gap of unknown length
9522 11052: contig of 1531 bp in length
11053 11152: gap of unknown length
11153 12497: contig of 1345 bp in length
12498 12597: gap of unknown length
12598 14043: contig of 1446 bp in length
14044 14143: gap of unknown length
14144 15764: contig of 1621 bp in length
15765 15864: gap of unknown length
15865 17363: contig of 1499 bp in length
17364 17463: gap of unknown length
17464 19002: contig of 1539 bp in length
19003 19102: gap of unknown length
19103 20200: contig of 1098 bp in length
20201 20300: gap of unknown length
20301 22038: contig of 1738 bp in length
22039 22138: gap of unknown length
22139 23176: contig of 1038 bp in length
23177 23276: gap of unknown length
23277 25200: contig of 1924 bp in length
25201 25301: gap of unknown length
25301 26417: contig of 1117 bp in length
26418 26517: gap of unknown length
26519 28095: contig of 1578 bp in length
28096 28195: gap of unknown length
28196 30110: contig of 1915 bp in length
30111 30210: gap of unknown length
30211 31622: contig of 1411 bp in length
31623 31721: gap of unknown length
31722 33428: contig of 1707 bp in length
33429 33528: gap of unknown length
33529 34953: contig of 1425 bp in length
34954 35053: gap of unknown length

35054 36516: contig of 1463 bp in length
36517 36616: gap of unknown length
36617 38338: contig of 1722 bp in length
38339 38438: gap of unknown length
38439 39587: contig of 1149 bp in length
39588 39687: gap of unknown length
39688 41558: contig of 1871 bp in length
41559 41558: gap of unknown length
41559 43328: contig of 1670 bp in length
43329 44329: gap of unknown length
44330 44703: contig of 1275 bp in length
44704 44804: gap of unknown length
44805 46421: contig of 1617 bp in length
46422 46520: gap of unknown length
46521 47825: contig of 1305 bp in length
47826 47925: gap of unknown length
47926 49770: contig of 1845 bp in length
49771 49870: gap of unknown length
49871 51415: contig of 1545 bp in length
51416 51515: gap of unknown length
51516 53139: contig of 1624 bp in length
53140 53239: gap of unknown length
53240 55131: contig of 1891 bp in length
55131 55230: gap of unknown length
55231 57213: contig of 1983 bp in length
57214 57313: gap of unknown length
57314 58882: contig of 1669 bp in length
58883 59082: gap of unknown length
59083 60935: contig of 1853 bp in length
60936 61035: gap of unknown length
61036 63349: contig of 2314 bp in length
63350 63450: gap of unknown length
63451 65305: contig of 1856 bp in length
65306 65405: gap of unknown length
65406 67518: contig of 2113 bp in length
67519 69828: gap of unknown length
69829 71369: contig of 2210 bp in length
71370 71469: gap of unknown length
71470 73263: contig of 1794 bp in length
73264 75861: contig of 2498 bp in length
75862 75961: gap of unknown length
75962 78826: contig of 2865 bp in length
78827 78926: gap of unknown length
78927 81105: contig of 2179 bp in length
81106 81205: gap of unknown length
81206 83738: contig of 2533 bp in length
83739 83838: gap of unknown length
83839 85845: contig of 2007 bp in length
85846 85945: gap of unknown length
85946 87480: contig of 1535 bp in length
87481 87580: gap of unknown length
87581 91423: contig of 3843 bp in length
91424 91523: gap of unknown length
91524 93192: contig of 1669 bp in length
93193 93292: gap of unknown length
93293 96456: contig of 3164 bp in length
96457 96556: gap of unknown length
96557 98470: contig of 1914 bp in length

Query Match

Best Local Similarity 75.0%; Score 474.6; DB 2; Length 154891;

Best Local Similarity 75.0%; Pred. No. 6.5e-69;

Matches 607; Conservative 0; Mismatches 199; Indels 3; Gaps 1;

Oy 1 ATCACTACAGGCGCATGAGCTGCGAGACAGGTGCTCCGCGCTTGGCG 60
||||| ||||||| ||| ||||||| ||||||| ||||| |||||
Db 19928 ATCACTACAGGCGCATGCGAGACAGGTGCTCCGCGCTTGGCGTA 19869
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||
Oy 61 CGTACCACAGCGCGAGCGACGATGAGGCGATGATGCTGACATTC 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||
Db 19868 CGTACCAGTCCGCGAGCGACGACGATGAGGCGATGATGCTGACATTC 19809

QY 121 CGCTGGAACGTGATCATTTGCTGTGTAGAGACGACACCTATGCGCGCAATGCGCAG 180
 19808 CAATGGAACGTGATGTTGGTGTGTGTAGAGACGACGATTAAGCGCGCAATGCGCAG 19749
 QY 181 CTGCTTGGGAGGCGGTGCGC---CGCGGACATCTGCATGCGCTTCCAGAGAGCGCTG 237
 19748 CTGTTAGCGACGCTGTGACAAAGAGAGACATCTGCATGCGCTTCCAGAGAGGTTCTG 19689
 QY 238 CCCACACTGAGCGCCACACGACATGAGCTGAGAGAGCGCAGCGCGCTGAGACAT 297
 19688 CCCATACCTGAGTCCAGCCAGGTCATGAGTCCGAGAGAGAGACACATGAGACATC 19629
 QY 298 GTGACAAAGCTGACAGAGACAGACGCGCGCTGCTGCTGCTTCCGCGCAGCTGAC 357
 19628 CTGACAAAGCTGCGCGGAGACCTGCGCGCGCTGCTGCTGCTTCCGCGCAGCTGAGC 19569
 QY 358 CTGTACACCTCTTCAATGAGTGTGCTGCGCAGACATCTTCCAGCGCGCGCTGAGTCCG 417
 19568 CTGTATAGCTCTTCTTCAAGAGGCTGCTGCGCAGACATCTTCCAGCGCGCTGAGTCCG 19509
 QY 418 TCCGAGCTGCGGCGCATGACGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 477
 19508 TCTGAGTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 19449
 QY 478 ACCCTCTGCGGCGCATGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 537
 19448 ACTTCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 19389
 QY 538 TGGGCGCGCAGACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 597
 19388 CGCGGCGCAGACG 19329
 QY 598 CAGAGTGCACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 657
 19328 CAGAGTGCACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 19269
 QY 658 GGGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 717
 19268 GGGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 19209
 QY 718 AGCTCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 777
 19208 AGCTCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 19149
 QY 778 CTGAGGAGATCTGGAAGTCAACTTCAC 806
 19148 GGTCCGAACTCGCAGGGGCTGACTCCAC 19120
 RESULT 12
 AFI27389 2520 bp mRNA linear ROD 04-MAR-1999
 LOCUS AF127389
 DEFINITION Rattus norvegicus putative taste receptor TR1 mRNA, partial cds.
 ACCESSION AF127389
 VERSION AF127389.1 GI:4337085
 KEYWORDS
 SOURCE Rattus norvegicus.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 2520)
 Hoon,M.A., Adler,E., Lindemeier,J., Battey,J.F., Ryba,N.J. and
 Zuker,C.S.
 TITLE Putative mammalian taste receptors: a class of taste-specific GPCRs
 with distinct topographic selectivity
 JOURNAL Cell 96 (4), 541-551 (1999)
 MEDLINE 99159821
 PUBMED 10052456
 REFERENCE 2 (bases 1 to 2520)
 Hoon,M.A., Adler,E., Lindemeier,J., Battey,J.F., Ryba,N.J.P. and
 Zuker,C.S.

TITLE Direct Submission
 JOURNAL Submitted (10-FEB-1999) Taste and Smell Unit, MIDCR, 10 Center
 Drive MSC 1188, Bethesda, MD 20892-1188, USA
 FEATURES
 source
 1.2520
 /organism="Rattus norvegicus"
 /strain="Wistar"
 /db_xref="taxon:10116"
 /cell_type="Circumvallate papilla"
 1..>2520
 /note="G. protein-coupled receptor"
 /codon_start=1
 /product="putative taste receptor TR1"
 /protein_id="AAD18069.1"
 /db_xref="GI:4337086"
 /translation="MLFMAHLLLSLQLYCMAFSCORTESPGRSLPDILLAGLS
 LHGDLQVHRLPIATSCDPDPSFNHGRVLRPMAPTYEINSSALLPITLIGELY
 DYCSSANYATLRLVALDQPHRIEORDRHHSSKYAFIFPDNTDHAATYTAALGP
 FLMPLYSVASSVLSAKRPFSLKTPSDRHOVEMVQLQSGWMI SLIGSYGD
 YQLGVQALEELAVPRGICVAFKDIIVFSAKRGDPRMOSMOMHLOARTVTVVSNR
 HLAFFRSVLANLNGKVVASDEMAISTYITSTYIGIGTGLVAVQOORVEGLK
 EFEESYRAVTAAPSACPEGSMCTNOLRECHETTRMPTLGAFSMAAYRYEAV
 VAVAGLHLLCTSEICRGPRVPMQLLOIYKXNLTHTVVAFPDNDGLGYDI
 IAMDNGEPTEITIGSASLSVHLDINRTKIOMGNKNQVYVSTCTDCLGHRRVY
 VSHHCECEVCENGTEFLNMSLHICQCGEEMAKRSTVCFPRVEFLAMHEPIS
 IYLIANTLLLLLVGTAGLFAMHRTVPVVRAGRLCLMLGSLVAGCSYSEFE
 PVPACILFQPLFSLGFAIFSLCIRSFQVLIIRFSTKVPFFRYTMQNHGACLFV
 IVSYVHLICLTWLVMPRTREYORFVGLTIECTVNSVGLFATPHNLLIS
 TFVSGTICKELPENNEAKCVFSLINRVSHVIAFTPMASITVGSYSLPVPVNLGLTT
 LSGFSGYRLPRCYVILCRPELNNEHROASIODYTRRCGTT"
 BASE COUNT 479 a 749 c 676 g 616 t
 ORIGIN
 Query Match 19.5%; Score 392.8; DB 10; Length 2520;
 Best local Similarity 51.4%; Pred. No. 3.2e-55;
 Matches 1046; Conservative 0; Mismatches 942; Indels 48; Gaps 4;
 QY 2 TCACCTACAGCGCCATGAGCGATGAGCGGACGACAAAGTGCCTTCCGCGCTTGTCTGC 61
 497 TCAGCTATGAGCGACAGCGAGTGTACTGATCCCAAGCCAAAGTTCCCTCTTCTCTTC 556
 QY 62 GTACACACCCAGCGCGCGACGACAGCTGAGGCGCATGAGCGATGCTCAGCTCC 121
 557 GTACCGTCCCACTGACCGGACGACGAGTGAGTCAATGAGCGCTGCAAGATTTC 616
 QY 122 GCTGGAACCTGATCATTTGCTGTGTAGAGACGACACCTATGCGCGCAATGCGCAGC 181
 617 GGTGGGTGTGTGATCTGCTGCTATTTGCGAGTATGAGGCGACGCTGCTGCTGAG 676
 QY 182 TGCTTGGCGAGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 241
 677 CGCTGAGAGAGCTGCG 734
 QY 242 CACTGACGCCACGACGACATGACGTCAGAGAGAGCGCGCGCGCGCGCGCGCGCGCG 301
 735 -----TTTCTGCGCGCGGTGCTGACCGGAGATGACAGAGCATGATGCG 778
 QY 302 ACAAGCTGACGACGACACGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
 779 AGCATCTGGCTGACGAGACGACGAGTGTGTGTGCTTCAACCGGACCTGGGCTA 838
 QY 362 ACCACTTCTTCAATGAGTGTGCTGCGCGACACTTTCACGGGCGCGGTGTGATGCTCG 421
 839 GAGTGTCTTCAAGTGTGCTGCTGCGCGCACTGATGCAAGAGTGTGTGCTGCTGAG 898
 QY 422 AGTCCGCGGCGCATGACGCGCGCTGCTGACACACTGACGAGCGCGGCGCACTTGGGACCT 481
 899 AAGACTGGGCGCATCTCAGCTACATCAACGCTGATGGGATCAAGGATTTGGGCGG 958
 QY 482 TCCTGGGCGATCACCATCAAGAGCGTCCCGCGGCTTCAAGTGTGCTGCGGAGTGG 541
 959 TGCTCGGTGTGCGCGCTGCTGACGACAGACAGTCCCTGCGGCTGAAGAGGTTGAGAGAGTCT 1018

QY	542	GCCACAGGCTGGGCGCCGACCCCTCAGCAGAGACACGACAGCTATACCTGC-----A	595
Db	1019	ATGTGAGGGGCTTGTAAACAGCTGCTCCACGCGCTTGGCCGAGGGGGTCTGTGGTACACACTA	1078
QY	596	ACCAAGAGTGGGACAACTGTGCTGAAAGGCCACTGTCTCTTCAACACACTTCAGAGGCT	655
Db	1079	ACCAAGCTGTGCGGGAGTGGCACACGTTCCACGACTGTAAACAGCCACGCTTGAAGCT	1138
QY	656	CTGGGGAGCGTGTGCTGTACAGGCTGTACTCTGCGGCTATCTGTGGCCATGGCCCTGC	715
Db	1139	TCTCCATGAGTGGCGGCTCTACAGAGTGTATGAGGCTGTATAGCTGTGGGCCACGGCCCTCC	1198
QY	716	ACAGCCTCTCGGCTGTGACAAAACACCTGCACCAAGAGGTGTCTACCCCTGGCAGC	775
Db	1199	ACCAAGCTCTGGGGATGTACTTCTGTGAATCTGTCTCCAGAGGCCCACTTACCCCTGGCAGC	1258
QY	776	TGCTTGGAGGAGATGTGAAGTCACTCACTCTCTCTGGACCAACAATCTTCTTCGAC	835
Db	1259	TTCTTTCACAGATATACAAAGTAAATTTCTTCTATAGAAATCTGTGGCAATTGTATG	1318
QY	836	CGCAAGGGAGCTGGCTCTCAGCTTGGAGATTGTCCAGTGGCAATGGGACCGAGCCAGA	895
Db	1319	ACAAGGGGAGACCTCTAGGTACTAGAGACATATGCGCTGGGAGCTGGAAATGACCTGAT	1378
QY	896	ATCCCTTCCAGAGCGCTGCCTCCTACTACCCCTGCAGCGACAGCTGAAGAACTAACA	955
Db	1379	GGACCTTGTGAGATCATTTGGCTCTGCGCTCACTGTCTCCAGTTATCTATGGCATAAATAGA	1438
QY	956	CATCTCT-----GCACACGGTCAACAACAACGATCCCTATGTCCAMTGTCTCCAGAGT	1009
Db	1439	CAAAATTCAGTGGCACGGGAGAAACAAATCAGTGGCTGTGTCAAGTGTACACAGGACT	1498
QY	1010	GCCAGTCAGGGCAAAAGAAAGAACCTGTGGGCTATCCAGCTTGTGCTGTCAAGCTACACG	1066
Db	1499	GTCGTGGAGGGCACACAGAGGTGTGTGGGGTCCACACACGTGTGAGTGTGTGTC	1558
QY	1070	ACTGCTTCCCGGACCTTCTCTAACCACACTGA-----TGCCTGA	1111
Db	1559	CTCTGAGAGCTGGGACCTTCTCTCAACATGATGAGCTTCACTTCGACAGCCCTTGTGAA	1618
QY	1112	ATAAGAGTGTGCTTACCAAGTAGAGACCTCCGCTTCAAGGGGACGAGTGTCTTCGG	1171
Db	1619	CAGAGAAATGGGACCTCAAGAGAGACCTACTGTCTCCACGACAGGGTGGATCTTGG	1678
QY	1172	AATGGATGAGGACACCCACACATGCTGTGGCCCTGTGGCCGCCCTTCCCTCCAGCA	1231
Db	1679	CTTGCAATGAACCATCTCTTTGTGTCTAATAGACACTAACACGCTATTTGTGTGTCTGC	1738
QY	1232	CCGTGGCCATCTGTGATATTTCTGAGGCACTTCCAGACACCCATAGTTCCTGGCGTG	1291
Db	1739	TGTTTGGGACGTGCTGGGCTGTGTGCTGTGGCATTTTCCACACCTGTAGTAGAGTCAAGT	1798
QY	1292	GGGGCCCCATGTGTTTCTCTATAGCTGACACATGTGTGTGGCAATCATGTGTGCCGG	1351
Db	1799	GGGGTAGGCTGTGTTCTCTATCTGTGGTTTCCCTGTGGCCGGAAGTTGCACACTCTCTATA	1858
QY	1352	TGTACGTGGGGCCGCCCAAGGCTCCACACCGCTGTGGCCAGGCGCCCTTCCCTCCCT	1411
Db	1859	GCTTCTTCCGGGAGCCACAGGTGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1918
QY	1412	GCTTCCAAATTTGCATCTCTGTATGTGCGGTCTTTTCCAGATTCGTGCGCCCTTCA	1471
Db	1919	GSTTTGCCATTTCTCTCTCCGTGCGTGCATATCGGCTCTTCAACAGTGCATCATCTTCA	1978
QY	1472	AGATGGCCAGCGCTTCCACGGGCTTACAGCTACTGGGTCCGCTACACAGGGGCCCTTACG	1531
Db	1979	AGTTTCTTACCAAGTCCCACTTCTATCCGTAACCTGGGCCCAAAACCATGTGTGAGGCTC	2038
QY	1532	TCTCATAGGCAATTTATACGAGTACTCAAAATGTGCTATTTGGAAATTTGGCAATGCGTGCAC	1591
Db	2039	TATTCGTATTTGTAGAGCTCCACAGGTCACTATTTGTCTACTCTGTCTCATATGTGCTTAATGT	2098
QY	1592	GGCCTCACTCCACCCCGTACTGACCCGATGACCCCAAGATTCACAAATTTGTCTCTCTTA	1651

Dd	2099	GGACCCCAAGACCACGAGGGGAATACGAGCGCTTCCCACATCTGTGTATTCTCGAGTGCA	2158
Qy	1652	ACCCCAACTACGGCAAAAGCCGTCGTTTCAACACACGAGCCTGGAGCTGTGCCTCAGTGG	1711
Dd	2159	CAGAGGTCAACTCTGTAGGCTTCTGTGGCTTTCACCCACACAATTTCTCTTCATCA	2218
Qy	1712	TGGGTTTCAGCTTCGCCATCATGSGCAAAGAGCTGCCACCACTACAGAGGCCAAGT	1771
Dd	2219	GTAACCTGTGTGACAGCTACCTGGGTAAAGAACTCCAGAGAATAATAGAACCAAT	2278
Qy	1772	TCATCACCTTCAGCATFAGACTTTCTATTTCACCTATCCGCTCTCCCTGTGACCTTCATGT	1831
Dd	2279	GHTGTCACCTTCAGCCTGCTCTCCACACTTCGATTCGATGCGCTTTCACACATGGCCA	2338
Qy	1832	CHGCCATACAGGGGGGTGGCGGACATCAGCGGAGACCTTGGTACCTGTGTCACACCTCC	1891
Dd	2339	GCATTTACAGGGGACAGCTACTGCTCGGGGTCATATGTCTGTGGACGGCTGCACACACTGA	2398
Qy	1892	TGGCCATCAGCCTGGGCTACTTCGGCCCCAAGTGCTACATGATCCTCTTCTACCGCGAGC	1951
Dd	2399	GGGGGGGGCTTAGCGGGTTACTTCCTCCCAAGTGTATGTCTGTGCGGCTCCAGAAC	2458
Qy	1952	GCAACACGCGCCCTACTTCAACAGCATGATTCAGGGCTACACATGAGAGGGAC	2007
Dd	2459	TCACACATACAGMACACTTTCAGGCGCTTCATCCAGAGCTACACGAGGGCGCGGC	2514
 RESULT 13 AX282894			
LOCUS	AX282894	2292 bp	DNA linear PAT 02-NOV-2001
DEFINITION	Sequence 3 from Patent WO0172842.		
ACCESSION	AX282894		
VERSION	AX282894.1	GI:16609869	
KEYWORDS	human.		
SOURCE	ORGANISM	human.	
		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
REFERENCE	1	Walke,D.W., Scoville,J., Donoho,G., Turner,C.A., Friedrich,G.C.,	
AUTHORS		Abidin,A., Zambrowicz,B. and Sands,A.T.	
JOURNAL		Novel human 7tm proteins and polynucleotides encoding the same	
		Patent: W0 0172842-A 3 04-OCT-2001;	
FEATURES		Lexicon Genetics Incorporated (US)	
	source	Location/Qualifiers	
		1..2292	
		/Organism="Homo sapiens"	
		/db_xref="taxon:9606"	
BASE COUNT	437 a 690 c 638 g 524 t	3 others	
ORIGIN			
Query Match	19.2%;	Score 386.4;	DB 6; Length 2292;
Best Local Similarity	51.1%;	Pred. No. 3.7e-54;	
Matches 1042; Conservative	1; Mismatches 947; Indels 48; Gaps 4;		
Qy	1	ATACCTACAGGGCATGAGGATGAGTGGCGAGACAGGTGCGCTTCCCGGCTTGTG	60
Dd	265	ATTAGCTTAGGGGGCGAGCGAGCGCTCACCGTGAAAGCGCGAGTATCCCTTCTG	324
Qy	61	CGTACACACCCAGCGCGAGCACCACAGCTGCAGGCGCATGTCAGCTGATGCTGCAC	120
Dd	325	CGCACCATTCCCAATGACAGATACAGAGGTGAGACACATGATGTCTGCTGCAGAA	384
Qy	121	CGCTGGAACTGATCATTTGTGTGTGAGACGACACACCTATGCGCGGACAAATGCG	180
Dd	385	GGGTGACCTGTGATCTCTGTGTGGACAGATGACGACTATGGGCACTGAGGGGTG	444
Qy	181	CTGCTTGGGAGAGGGGTGGCCCGGGGCGACATCTGCAATGCGCTTCAGAGAGCGT	240
Dd	445	GCACCTGGAAACAGGCGACCTGTGTAGGGAGATCTCATTTGTTTAAAGACATCAT	504
Qy	241	ACACTGAGCCCAACAGAACATGACGTACAGAGAGCGCCACGCGCTGTGACCAT	300

Db 505 -----TCTCTGCCAGAGTGGGCGATGAGAGATGACAGTCTCATG 546
 QY 301 GACAACTGACAGAGACACAGGGCGGCTGCTGTCTCTCCGCCACCTGACCC 360
 Db 547 CGCAGCTGAGCCAGGCGGGGCGCAGCTGTGTGTCTTCTTCCAGCGCGAGTGGCC 606
 QY 361 TACCACTTCTTCAATGAGTGTCTGCGCAGAACCTTCAAGGGCGCGGTGATGCTCC 420
 Db 607 AGGGTGTCTTTCAGTGCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 666
 QY 421 GAGTCTGGGCGATGACAGCCGCTGTCTGACACACCTTACAGAGCTGGGCGAC 480
 Db 667 GAAGCTGGGCGCTTCTTCCAGGACATCATCTGGGCTGGGCGATGACAGCCATTTGGGATG 726
 QY 481 TTTCTGGGCGATGACACCTTCCAGGCGATGCGGCGCTTCACTGATGCTTCCGAGTGG 540
 Db 727 GTCTTGGGCGATGACACCTTCCAGGCGATGCGGCGCTTCACTGATGCTTCCGAGTGG 786
 QY 541 GGGCCACAGGCTGGGCGCGCCCTTCCAGGCGATGCGGCGCTTCACTGATGCTTCCGAGTGG 594
 Db 787 TATGCCCGGCGACAGAGAGGCGCCCTTGGCGACAAAGGCGCTTCTGTGACAGGCG 846
 QY 595 AACCAAGAGTGGCAGCACTGCTGAAAGCCACTTGTCTTCAACACCATTTCTGAGGCTC 654
 Db 847 AATCAGCTCTGACAGAGATGCGCAAGCTTTCATRCACACAGATGCGCAAGCTCAAGGCC 906
 QY 655 TCTGGGAGGCGTGTCTGCTTACAGCTGTACTGCGGCTCATCTGCTGGGCGATGCGGCGCT 714
 Db 907 TTTCTCATGATGTCTGCTTCAACAGCATACCGGCTGTGTGTGTGTGTGTGTGTGTGTGT 966
 QY 715 CACAGCTCTCTGCTGTGTGCAAAAGCACTGACCAAGAGGCTGTCTACCCCTGGCAG 774
 Db 967 CACCACTCTCTGCGGCTGTGTGCTTGTGAGCTTGTTCAGAGGCGCGAGCTTCCCTGGCAG 1026
 QY 775 CTGCTTGAAGAGTCTGGAAGTCACTTCACTCTCTGACCAACCAATCTTCTTGAC 834
 Db 1027 CTTTGTGAGAGATCAACAAAGGTGATTTCTTCTTCAACAAAGACACTGTGGCTTTAAT 1086
 QY 835 CCGCAAGGAGGAGCTGTGCTGTGAGATGTGCAAGTGTGCAAGTGTGAGGCGGAGGAGC 894
 Db 1087 GACACAGAGATGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1146
 QY 895 AATCCTTCCAGAGCTGCTCTCTAC-----TACCCCTGACGAGCAGCTGAAGAAC 948
 Db 1147 TGGACCTTCAAGGCT 1206
 QY 949 ATCAAGACATCTCTGACACCCGTCACACACATCCCTATGTCTTCCATGTCTTCCAAAGAG 1008
 Db 1207 ACCAAATTCAGTGGACGAGAAAGACACAGGTGCTAAGTCTGTGTCTTCCAGGAGC 1266
 QY 1009 TGGCAGTCAAGGCGCAAAAGAAAGAGCTGTGGGATGCAAGCTGTGCTGTGATGCTGATG 1068
 Db 1267 TGTCTTGAAGGAGCAGCAGAGAGTGTGTAGGCTTTCATGATGCTTGTGTGTGTGTGTGT 1326
 QY 1069 GACTGCTTCCCGGCACTTCTCTCAACACACTGAA-----TGCCTG 1110
 Db 1327 CCTGTGGGCGCTGAGACTTCTCTCAACAGAGAGCTTCTTACAGATGACAGCTTGTGGG 1386
 QY 1111 AATAAGAGAGTGTCTTCAAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1170
 Db 1387 AAGAAAGAGT 1446
 QY 1171 GAATGAGT 1230
 Db 1447 GCTTGT 1306
 QY 1231 ACCCTGAGCATCTGT 1290
 Db 1507 CTGCTTGT 1566
 QY 1291 GGGGGGCGCATGT 1350

Db 1567 GGGGGGCGCATGT 1626
 QY 1351 GTGTAGT 1410
 Db 1627 GGTCTTGT 1686
 QY 1411 TGTCTTCAATTTGT 1470
 Db 1687 GGTTCACATCTTCTGT 1746
 QY 1471 AAGATGCGCAGCGCTTCCAGCGCTTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1530
 Db 1747 AAGTTTCCACCAAGGT 1806
 QY 1531 GTCTATGAGCATTTTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1590
 Db 1807 CTGT 1866
 QY 1591 CGGCTGT 1650
 Db 1867 TGGACCCCACTGCTGT 1926
 QY 1651 AACCCCACTACCGCAACAGCTGT 1710
 Db 1927 ACAGAGACCAACTCTCTGT 1986
 QY 1711 GTGGT 1770
 Db 1987 AGTGCCTTGT 2046
 QY 1771 TTTATACCTTGT 1830
 Db 2047 TGT 2106
 QY 1831 TGT 1890
 Db 2107 AGCTGT 2166
 QY 1891 CTGGCCTATGAGCTGT 1950
 Db 2167 AGCAGGCGCTTGT 2226
 QY 1951 CGCAGACGCGCGCTTGT 2008
 Db 2227 CTCACAGACAGAGAGCTTCCAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2284
 RESULT 14
 AX282892
 LOCUS AX282892 2526 bp DNA linear PAT 02-NOV-2001
 DEFINITION Sequence 1 from Patent WO0172842.
 ACCESSION AX282892
 VERSION AX282892.1 GI:16609868
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Walke,D.W., Scoville,J., Donoho,G., Turner,C.A., Friedrich,G.C.,
 Abuln,A., Zamrowicz,B. and Sands,A.T.
 TITLE Novel human 7em proteins and polynucleotides encoding the same
 JOURNAL Patent: WO 0172842-A 1 04-OCT-2001;
 Lexicon Genetics Incorporated (US)
 FEATURES
 source 1..2526
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 469 a 770 c 697 g 587 t 3 others
 ORIGIN
 Query Match 19.2%; Score 386.4; DB 6; Length 2526;
 Best Local Similarity 51.1%; Pred. No. 3.7e-54;

Matches 1042; Conservative 1; Mismatches 947; Indels 48; Gaps 4;

```
QY 1 ATCACTTACAGGCGCATGCGATGAGTCGCGAGACAGGTGCGTCCCGGCTTTCG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 ATTAGCTATGCGCGCCAGCGAGCGAGCGTCAAGGTGACGCGATATCCCTTTCG 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 CGTACCAACCGAGCGCGCGCCAGCGAGTCAAGGCGATGCGATGATGCGACCTTC 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 559 CGACCATATCCCATATGACAAATACAGTGAACCATGCTGCTGCTGCAAGATTC 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 CGCTGGAATGATCATTTGCTGTGTGAGCAGCGACCATATGAGCGCCACATATGCG 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 GGGTGAACCTGGATCTCTGTGTGAGCAGATGACATATGAGGAGCGATGAGGAG 678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 CTGCTGGGAGCGCGTGGCGCGCGAGATCTGATGCGCTTCAGAGACGCTGCC 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 GCACGTGAAGAACCGCGCATGCTGAGGGGATCTGATGCTTCAAGACATCATAGCCC 738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 ACACGTGACCGCCAAACAGAACATGATGAGAGCGCGCGCGCTGGTGAACATTTG 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 -----TTCTGTGCGCCAGGTGGCGGATGAGAGATGACATGCTCATG 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 GACAACTGACAGACAGACAGCGCGCTGTGTGCTGTCTCGCCGACCTGACCTG 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 781 CGCACACTGCGCCAGCGCGCGCGCGCGCTGTGTGCTGTCTCGCCGACCTGACCTG 840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 TACCACTTCTCATATGAGTGTGCGCGCGCGCGCGCGCTGTGTGCTGTCTCG 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 841 AAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 GATGCTGGGCGCATGACCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 901 GAAAGCTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 481 TTCTGTGGGATGACATGACAGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGT 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 961 GTGCTGTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 541 GGGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 594
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1021 TATGCGCGGCGAGACAGAGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 595 AACCAAGATGCGACATGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 654
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1081 AATCACTGTGACAGAAATGCGCAAGCTTTCATGACACATGATGCGCAAGCTT 1140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 655 TCTGGGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 714
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1141 TTCTCATGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 715 CACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1201 CACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 775 CTGCTGTGAGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 834
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1261 CTTTTGGACAGATTCACAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 835 CCGCAAGGAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 894
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1321 GACAAAGAGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 895 AATCCCTTCCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 948
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1381 TGGACCTTACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 949 ATCAAGACATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1441 ACCAAAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1009 TGGCACTGAGGCAAAAGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1068
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1501 TGTCTTGAAGGACGACGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 1069 GACTGCTTCCCGGCGACCTTCTCAACGACACTGAA-----TGCCTG 1110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1561 CCTGTGGGCGGTGGGAGCTTCTCAAGAGAGTCAAGCTTCAAGATGACGCTGTGG 1620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1111 AATTAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1621 AATTAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1680
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1171 GAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1681 GCTTTGCGTAGACACACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1740
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1231 ACCGTGGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1741 CTGCTTGGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1800
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1291 GGGGGCGCGCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1801 GGGGGCGCGCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1351 GTGTACGTGGGCGCGCGCAAGTCTCACCTGTGTGTGTGTGTGTGTGTGTGT 1410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1861 GCGTCTTGTGGGAAACCAAGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1920
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1411 TGTTCACAAATTTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1921 GGTTCACCAATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1980
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1471 AAGATGGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1981 AAGTTTCCACCAAGTACTTACATTTACACGCTGTGTGTGTGTGTGTGTGTGT 2040
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1531 GTCTGTATGAGCAATTTATACAGGTACTCAAAATGTGTGTGTGTGTGTGTGT 1590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2041 CTGTGTGTATGATACGTACAGGCGCGAGCTGTGTGTGTGTGTGTGTGTGT 2100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1591 CGGCGTCACTCCACCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2101 TGGACCCGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1651 AACCCCACTACCGCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2161 ACAGAGACCAACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1711 GTGGGTTTACGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1770
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2221 AGTGCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1771 TTATACACCTGACATGACCTTTATTTACCTCATCCGTCTCCCTGTGACCT 1830
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2281 TGTGTACCTTACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1831 TGTGCTTACAGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1890
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2341 AAGCTGTACAGCGCAAGTACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 2400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1891 CTGGGCAATGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1950
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2401 ACGAGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1951 CGCAACAGCGCGCTTACTTCAACAGCATGATCCAGGCTTACACCATGTAG 2008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2461 CTCAACAGCAGACAGCTTCAAGGCTTCAATTCAGGACTACAGGCGCTGG 2518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 15
AX282900
LOCUS AX282900 2951 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 9 from Patent WO0172842.
ACCESSION AX282900
VERSION AX282900.1 GI:16609872
KEYWORDS

Db 2572 AGCGTCTACGAGCGGCAAGTACCTGCGGCCAACATGATGGCTGGGCTGAGCAGCCTG 2631
OY 1891 CTGGCCATCAGCCTTGCGCTACTTCGGCCCCAAGTCCTACATGATCCTCTTACCCGGAG 1950
Db 2632 AGCAGCGGCTTCGATGGGTAATTTCTGCTAGTGTGATCTCTGCGCCCGCAGAC 2691
OY 1951 CGCAACAGCGCCGCTACTTCAACAGCATGATCCAGGGCTACCCATGAGAGGAGCT 2008
Db 2692 CTCACAGCAGCAGACACTTCCAGGCTTCATTTCAGAGACTACAGAGGGCGCTGCGGCT 2749

Search completed: May 23, 2003, 16:36:36
Job time : 5982.99 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 11:20:10 ; Search time 420.125 Seconds
(without alignments)
10774.220 Million cell updates/sec

Title: US-09-927-315-12

Perfect score: 2010
Sequence: 1 ataccctacgcgcgcacacag.....acacacatgagagggactag 2010

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2010	100.0	2010	AAZ58965 Human GPCR-B4 poly
2	1232.8	61.3	2993	21 AAZ58963 Rat GPCR-B4 poly
3	1212.6	60.3	2532	21 AAZ58963 Mouse GPCR-B4 poly
4	392.8	19.5	2771	21 AAZ50743 Rat sensory transd
5	386.4	19.2	2292	22 AAD19502 Human novel G-prot
6	386.4	19.2	2526	22 AAD19501 Human novel G-prot
7	386.4	19.2	2951	22 AAD19505 Human novel G-prot
8	385.2	19.2	2759	24 ABL5953 Human G-protein co
9	376.6	18.7	2333	21 AAZ50745 Human sensory tran

10	368.8	18.3	2526	22 AAD17516
11	359.2	17.9	2579	21 AAZ50744
12	258	12.8	697	22 AAZ56935
13	258	12.8	697	24 ABAK81713
14	218.6	10.9	2559	24 ABAK47345
15	218.6	10.9	3458	24 AAL38462
16	217	10.8	259	24 AAS97395
17	217	10.8	3489	22 AAS14574
18	216.2	10.8	2553	22 AAD17509
19	187	9.3	4131	18 AAT89290
20	187	9.3	4134	24 ABL5964
21	187	9.3	4134	24 ABL14910
22	184.2	9.2	1101	22 AAD19503
23	182.2	9.1	8191	22 AAD17515
24	180.2	9.0	2729	24 ABL5952
25	172.6	8.6	5275	19 AAV6962
26	172.6	8.6	5275	19 AAT95857
27	172.6	8.6	5275	20 AAV25053
28	172.6	8.6	5275	20 AAV82483
29	172.6	8.6	5275	21 AAZ89296
30	172.6	8.6	5275	24 AAT72120
31	167	8.3	2703	21 AAZ50617
32	167	8.3	2922	21 AAZ50616
33	167	8.3	3003	21 AAZ50615
34	167	8.3	3234	20 AAX90922
35	167	8.3	3234	20 AAZ31046
36	167	8.3	3234	21 AAZ51398
37	167	8.3	3234	24 AAL43269
38	167	8.3	3809	17 AAT61382
39	167	8.3	3809	19 AAV26964
40	167	8.3	3809	19 AAT95859
41	167	8.3	3809	20 AAZ25055
42	167	8.3	3809	20 AAV82485
43	167	8.3	3809	21 AAZ89298
44	167	8.3	3809	22 AAT86094
45	167	8.3	3809	24 AAT72122

ALIGNMENTS

RESULT 1	AAZ58965	AAZ58965 standard; DNA; 2010 BP.
AC	AAZ58965;	
DT	08-MAY-2000	(first entry)
DE	Human GPCR-B4 polypeptide encoding DNA.	
KW	Sensory transduction G-protein coupled receptor; GPCR; GPCR-B4; human; taste transduction pathway; taste receptor; foliate; fungiform; food; circumvallate; taste signaling; pharmaceutical; ds.	
OS	Homo sapiens.	
FT	key	Location/Qualifiers
FT	CDS	1..2010
FT		/tag= "a
FT		/product= "GPCR-B4"
XX	WO200006593-A1.	
XX	PD	10-FEB-2000.
XX	XX	27-JUL-1999;
XX	PF	99WO-US17104.
XX	XX	28-JUL-1998;
XX	PR	98US-0095464.
XX	XX	17-DEC-1998;
XX	XX	98US-0112747.
XX	PA	(REGC) UNIV CALIFORNIA.
XX	XX	

PI Zuker CS, Adler JE, Lindemeier J;
XX WPI: 2000-195257/17.
DR P-PSDB; AAT77558.
XX
PT New isolated sensory transduction G-protein coupled receptor, useful
PT for developing products for use in studying and modulating the taste
PT transduction pathway and for generating taste topographic maps -
XX
PS Claim 5; Page 73-74; 76pp; English.
XX
CC The invention provides nucleic acids encoding rat, mouse and human
CC sensory transduction G-protein coupled receptor (GPCR) polypeptides. The
CC GPCR polypeptides are components of the taste transduction pathway. The
CC nucleic acids can be used to identify taste cells and as tools for the
CC generation of taste topographic maps that elucidate the relationship
CC between the taste cells of the tongue and taste sensory neurons leading
CC to taste centers in the brain. GPCR-B4 is useful as a nucleic acid probe
CC for identifying subpopulations of taste receptor cells such as foliate,
CC fungiform, and circumvallate taste receptor cells. The polypeptides can
CC be used for identifying compounds that modulate sensory signaling in
CC sensory cells. Such modulators of taste transduction are useful for
CC pharmacological and genetic modulation of taste signaling pathways. These
CC modulatory compounds can then be used in the food and pharmaceutical
CC industries to customize taste. The present sequence represents a DNA
CC encoding a human GPCR-B4 polypeptide.
XX
SQ Sequence 2010 BP; 387 A; 691 C; 515 G; 417 T; 0 other;
Query Match 100.0%; Score 2010; DB 21; Length 2010;
Best Local Similarly 100.0%; Pred. No. 0;
Matches 2010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 ATCACTACAGCGGCATCAGGATGAGTGGAGACAAGTGGCTTCCGGCTTGGTG 60
DB 1 ATCACTACAGCGGCATCAGGATGAGTGGAGACAAGTGGCTTCCGGCTTGGTG 60
QY 61 CGTACACACCCAGCGCGACACAGTGCAGGCGCATGTGTCAGTCTGACACTTC 120
DB 61 CGTACACACCCAGCGCGACACAGTGCAGGCGCATGTGTCAGTCTGACACTTC 120
QY 121 CGGTGGAACGTGATCTGCTGTGTGAGCAGCAGCACTTGGCCGCGACAAATGGCCAG 180
DB 121 CGGTGGAACGTGATCTGCTGTGTGAGCAGCAGCACTTGGCCGCGACAAATGGCCAG 180
QY 181 CTGCTTGGCGAGCGCGTGGCCGCGGACATCTGATCGCTTCCAGAGACGCTGCC 240
DB 181 CTGCTTGGCGAGCGCGTGGCCGCGGACATCTGATCGCTTCCAGAGACGCTGCC 240
QY 241 ACACTGACGCCAACAAGACATGATGAGAGAGCGCGCGCTGTGATGACATTGTG 300
DB 241 ACACTGACGCCAACAAGACATGATGAGAGAGCGCGCGCTGTGATGACATTGTG 300
QY 301 GACAGCTGACAGCAGACAGCGCGCTGTGCTGTCTGCTGCGCCGACCTGACCTG 360
DB 301 GACAGCTGACAGCAGACAGCGCGCTGTGCTGTCTGCTGCGCCGACCTGACCTG 360
QY 361 TACCACTTCTCAATGAGTGTGCTGCGCAGAACTTCAAGGGCGCGTGTGATGCGCTCC 420
DB 361 TACCACTTCTCAATGAGTGTGCTGCGCAGAACTTCAAGGGCGCGTGTGATGCGCTCC 420
QY 421 GAGTCTGAGGCGATGACCGCGTCTGACAACTTCAAGAGTGGGCGCAATTGGGACCC 480
DB 421 GAGTCTGAGGCGATGACCGCGTCTGACAACTTCAAGAGTGGGCGCAATTGGGACCC 480
QY 481 TTCTGTGGGCGATCAGCAATCCAGAGCGTCCCAATCCGGGCTTCAATGATCCGGAATGG 540
DB 481 TTCTGTGGGCGATCAGCAATCCAGAGCGTCCCAATCCGGGCTTCAATGATCCGGAATGG 540
QY 541 GGCCACACAGGCTGGGCGCGCAACCCCTCAGAGAGCGACGAGACTTATCTGACACGAG 600
DB 541 GGCCACACAGGCTGGGCGCGCAACCCCTCAGAGAGCGACGAGACTTATCTGACACGAG 600

QY 601 GAGTGGCAACAACGCTGTGAAGCGCACCTGTGCTTCAACACATTTCAAGCTCTGGG 660
DB 601 GAGTGGCAACAACGCTGTGAAGCGCACCTGTGCTTCAACACATTTCAAGCTCTGGG 660
QY 661 GAGGCTGTGCTTCAACAGCGTGTACTGCGGTCTATGCTGTGGCCATGCGCTGCACAGC 720
DB 661 GAGGCTGTGCTTCAACAGCGTGTACTGCGGTCTATGCTGTGGCCATGCGCTGCACAGC 720
QY 721 CTCTGGGCTGTGCAAAAGACCTGACCAAGGGGTGTCTAAGCCCTGGACGCTCTT 780
DB 721 CTCTGGGCTGTGCAAAAGACCTGACCAAGGGGTGTCTAAGCCCTGGACGCTCTT 780
QY 781 GAGGAGATCTGGAAGTCAACCTTCACTCTCTGAGCACCACAAATCTTTCAGACCCCAA 840
DB 781 GAGGAGATCTGGAAGTCAACCTTCACTCTCTGAGCACCACAAATCTTTCAGACCCCAA 840
QY 841 GGGGAGCTGGCTGTGCACTTGGAGATTGTCCAGTGGCAATGGGACCGGACCGAATGCC 900
DB 841 GGGGAGCTGGCTGTGCACTTGGAGATTGTCCAGTGGCAATGGGACCGGACCGAATGCC 900
QY 901 TTCCAGAGCGTGGCGCTCTACTACCTACCCCTGACAGGACAGCTGAAACATCAAGACTCT 960
DB 901 TTCCAGAGCGTGGCGCTCTACTACCTACCCCTGACAGGACAGCTGAAACATCAAGACTCT 960
QY 961 CTGCACACCGTCAACAACAGATCCCTATGTCCATGTGTTCCAAAGAGTGCAGTCAAGG 1020
DB 961 CTGCACACCGTCAACAACAGATCCCTATGTCCATGTGTTCCAAAGAGTGCAGTCAAGG 1020
QY 1021 CAAAAGAAAGAGCTGTGGGATCCAGCTGCTGCTTCCAGATGACATGCTGCTTCC 1080
DB 1021 CAAAAGAAAGAGCTGTGGGATCCAGCTGCTGCTTCCAGATGACATGCTGCTTCC 1080
QY 1081 GGCACCTTCTCAACACACATGATGCGCGAATAGAGTGGTCTTCAAGAGTGAAGC 1140
DB 1081 GGCACCTTCTCAACACACATGATGCGCGAATAGAGTGGTCTTCAAGAGTGAAGC 1140
QY 1141 TCCTGCTTCAAGCGGCGAGTGTCTTCTGTAATGAGCATGAGGACCCACCATGCTGTG 1200
DB 1141 TCCTGCTTCAAGCGGCGAGTGTCTTCTGTAATGAGCATGAGGACCCACCATGCTGTG 1200
QY 1201 GCGCTGTGGCGCGCTTCCAGACACCTGCGCATCTGCGTGTGATGATTTGAGG 1260
DB 1201 GCGCTGTGGCGCGCTTCCAGACACCTGCGCATCTGCGTGTGATGATTTGAGG 1260
QY 1261 CACTTCCAGACACCATAGTGTGCTGCGTGGGCGCCCATGTGCTTCCGATGCTGACA 1320
DB 1261 CACTTCCAGACACCATAGTGTGCTGCGTGGGCGCCCATGTGCTTCCGATGCTGACA 1320
QY 1321 CTGCTGTGTGCAATGATGTGTGCTCGGTGTGATGAGTGGGCGCCCAAGTCTCCACC 1380
DB 1321 CTGCTGTGTGCAATGATGTGTGCTCGGTGTGATGAGTGGGCGCCCAAGTCTCCACC 1380
QY 1381 TGCTTGTGCGGCGACGCGCTTTCCTTCCCTGTGCTTCAATTTGCAATCTCTGTATGCC 1440
DB 1381 TGCTTGTGCGGCGACGCGCTTTCCTTCCCTGTGCTTCAATTTGCAATCTCTGTATGCC 1440
QY 1441 GTGCGTCTTTCAGATGCTGTGCGCTTCAAGATGCGCAGCGCTTCCACGCGCTTAC 1500
DB 1441 GTGCGTCTTTCAGATGCTGTGCGCTTCAAGATGCGCAGCGCTTCCACGCGCTTAC 1500
QY 1501 AGCTACTGGGTGCGCTTACAGAGGGGCGCTTACGTCTATGAGCATTTATACAGGTACTCAA 1560
DB 1501 AGCTACTGGGTGCGCTTACAGAGGGGCGCTTACGTCTATGAGCATTTATACAGGTACTCAA 1560
QY 1561 ATGCTCATTTGTGTAATGAGATGCTGCGACGCGCTTACATCCACCCCGCTTACGACCC 1620
DB 1561 ATGCTCATTTGTGTAATGAGATGCTGCGACGCGCTTACATCCACCCCGCTTACGACCC 1620
QY 1621 GATGACCCCAAGATCAATTTGCTCTGTACCCCAACATACGCAACAGCAGCTGCTGTTCC 1680
DB 1621 GATGACCCCAAGATCAATTTGCTCTGTACCCCAACATACGCAACAGCAGCTGCTGTTCC 1680
QY 1681 AACACAGCCTTGAGACCTGTCTCTGATGTGTGAGTGGGTTTCAAGTTCGCTTACATGGGCGAAA 1740

```

|||||
Db 1681 AACACACCTGAGACTCTGCTGCTCACTGGTGGTTTCACTTGGCTTACATGGGCAA 1740
Qy 1741 GAGTGGCCCACTACAGAGGAGCAAGTTCATCACCCTCAGCATGACCTTCTATTTC 1800
Db 1741 GAGTGGCCCACTACAGAGGAGCAAGTTCATCACCCTCAGCATGACCTTCTATTTC 1800
Qy 1801 ACCTATCGCTCTCCCTGACACCTTCATGTCCTTACAGGGGGTGTGTACACATC 1860
Db 1801 ACCTATCGCTCTCCCTGACACCTTCATGTCCTTACAGGGGGTGTGTACACATC 1860
Qy 1861 GTGAGACTTGTGCTACTGTGCTCAACCTCGCCATCAGCTGGCTACTTGGCCCC 1920
Db 1861 GTGAGACTTGTGCTACTGTGCTCAACCTCGCCATCAGCTGGCTACTTGGCCCC 1920
Qy 1921 AAGTCTACATGATCTCTTCTTACCCGAGGCAACAGCCCGCTACTTCAACAGCATG 1980
Db 1921 AAGTCTACATGATCTCTTCTTACCCGAGGCAACAGCCCGCTACTTCAACAGCATG 1980
Qy 1981 ATCCAGGCTACACCATGAGAGGAGCTAG 2010
Db 1981 ATCCAGGCTACACCATGAGAGGAGCTAG 2010

```

RESULT 2
AAZ58963
ID AAZ58963 standard; DNA; 2993 BP.

AAZ58963;
08-MAY-2000 (first entry)

Rat GPCR-B4 polypeptide encoding DNA.

Sensory transduction G-protein coupled receptor; GPCR; GPCR-B4; rat;
taste transduction pathway; taste receptor; foliate; fungiform; food;
circumvallate; taste signaling; pharmaceutical; ds.

Rattus sp.

Location/Qualifiers
13..2544
/*tag= a
/product= "GPCR-B4"

W0200006593-A1.

10-FEB-2000.

27-JUL-1999; 99W0-US17104.

28-JUL-1998; 98US-0095464.

17-DEC-1998; 98US-0112747.

(REGC) UNIV CALIFORNIA.

Zuker CS, Adler JE, Lindemeier J;

WPI. 2000-195257/17.

P-PSDB; AAY77556.

New isolated sensory transduction G-protein coupled receptor, useful
for developing products for use in studying and modulating the taste
transduction pathway and for generating taste topographic maps -

Claim 5, Page 70-71; 76pp; English.

The invention provides nucleic acids encoding rat, mouse and human
sensory transduction G-protein coupled receptor (GPCR) polypeptides. The
GPCR polypeptides are components of the taste transduction pathway. The
nucleic acids can be used to identify taste cells and as tools for the
generation of taste topographic maps that elucidate the relationship
between the taste cells of the tongue and taste sensory neurons leading

to taste centers in the brain. GPCR-B4 is useful as a nucleic acid probe
for identifying subpopulations of taste receptor cells such as foliate,
fungiform, and circumvallate taste receptor cells. The polypeptides can
be used for identifying compounds that modulate sensory signaling in
sensory cells. Such modulators of taste transduction are useful for
pharmacological and genetic modulation of taste signaling pathways. These
modulatory compounds can then be used in the food and pharmaceutical
industries to customize taste. The present sequence represents a DNA
encoding a rat GPCR-B4 polypeptide.

Sequence 2993 BP; 619 A; 962 C; 729 G; 683 T; 0 other;

Query Match 61.3%; Score 1232.8; DB 21; Length 2993;

Best Local Similarity 76.9%; Pred. No. 5,6e-249; Mismatches 444; Indels 30; Gaps 4;

```

Matches 1568; Conservative 0; Mismatches 444; Indels 30; Gaps 4;
Qy 1 ATCACCACAGCCGATCAGGATGAGCTGCGAGACAGAGTGGCTTCCGCGCTTGTG 60
Db 505 ATCACCACAGCCGATCAGGATGAGCTGCGAGACAGAGTGGCTTCCGCGCTTGTG 564
Qy 61 GTTACACACCCGAGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120
Db 565 CGCAGAGTCCGAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 624
Qy 121 CGCTGGAACCTGATCATTTGTGCTGTGAGCAGCAGCAGCAGCAGCAGCAGCAG 180
Db 625 CATGGAACCTGATCATTTGTGCTGTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 684
Qy 181 CTGCTTGGCGAGCGCGTGGCC---CGGCGGACATCTGTCATGCGCTTCCAGAGACGCTG 237
Db 685 CTGTTGAGCGCAGCGTGTGACCAAAACGAGCAGCATGTGCTTCCAGAGAGTTCG 744
Qy 238 CCCACATCCAGCCCAACAGAAACATGAGTCAAGAGAGAGAGAGAGAGAGAGAG 297
Db 745 CCCATACCTGATCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 804
Qy 298 GTGACAAACCTCAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 357
Db 805 CTGGAACAGCTCGGCGGAGCCTCGGCGCGCTGCTGTGTGTGTCTCGGCGCAGCTGAGC 864
Qy 358 CTGTACCACTTCTTCATGAGTGTGCTGCGCAGAACTTCAGGCGCGCTGTGAGATGCC 417
Db 865 CTGTATAGCTTCTTTCACAGAGTGTGCTGCGCAGAACTTCAGGCGCGCTGTGAGATGCC 924
Qy 418 TCGAGTCTGTGGCCATGACCCGCGCTGCAAACTTCAGGAGCTGGGCACTTGGGCG 477
Db 925 TCTGAGTCTGTGGCCATGACCCGCGCTGCAAACTTCAGGAGCTGGGCACTTGGGCG 984
Qy 478 ACCCTCTGCGGATCAGCATCCAGAGCTGCCATCCCGGCTTCAGTGAATTCGCGAG 537
Db 985 ACTTTTCTGGGGGTACCATCCAGAGAGGTGTCATCCCTGTGAGTCCGAGTGG 1044
Qy 538 TGGGGCCCAAGAGCTGGGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 597
Db 1045 CGCGGTGACAAAGCAGGAGTATCCGCTTACACAGCAACCACTTGGGAGCAGCTCAAC 1104
Qy 598 CAGGAGTGGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCT 657
Db 1105 CAGGAGTGGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCT 1164
Qy 658 GGGAGAGCTGTGCTGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 717
Db 1165 GGGAGAGCTGTGCTGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1224
Qy 718 AGCTCTCTGGCTGTGACAAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 777
Db 1225 AGACTCTCTGGCTGTGACAAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1284
Qy 778 CTGAGAGATGTGGAAGGCAACTTCACCTTCTGAGCAGCAGCAGCAGCAGCAGC 837
Db 1285 CTCAGGAGATGTGGAAGGCAACTTCACCTTCTGAGCAGCAGCAGCAGCAGCAGC 1344
Qy 838 CAAAGGAGCTGCTGTGACCTTGAGATGTTCAGTGGCAATGGAGACCGAGCCAGAA 897

```

Db	1345	CAAGGGGACATGGCGGATGCTCTGGACATCATCCATGGACATGGGACCTGAGCCAGAT	11404
Qy	898	CCCTTCCAGAGCGTGGCTCTCTACATACCCCTGACGACAGCTGAAGAACAT--CAAG	954
Db	1405	CCCTTCCAAAGCATGGCTCTCTATCTCTCCACGACGACAGAGGCTAACCTACTTAAACAT	1464
Qy	955	ACATCTGCGACACGTCGACCAACAGATCCCATGTCATGTCATGTCGAAGGGGGCCAG	1014
Db	1465	GTGTCTGTGATACACCCCCCAACAGGTCCTCTGTCTCATGTGTTCAGAGGCTGCCAG	1554
Qy	1015	TCAGGCGAAAAGAAAGACCTGTGGGATCCACAGTCGCTGCTTGAGTGATGACATGC	1074
Db	1525	CCAGGGCAAAATGAAGATGCTGTGGGGCTCACCCCTTGTGCTCGAAGTGCTTGGAATGT	1584
Qy	1075	CTTCCCGGACACTTCTCTAACACACTGA-----TGCCCAAT	1113
Db	1585	ATGCGACGACATTCATCTCAACCCGCTCAGCAGATGATTTAACTGTGTCTGCGCCGGGT	1644
Qy	1114	AACGATGTCCTTACAGAGTGAAGCTCCCTGCTTCAAGCGGACAGCTGGTCTCTCGAA	1173
Db	1645	TCCATGTGTCTTCAAGAAAGCATGATGTGTTCCACAGGGGGGCTCAACTCTCTGGAG	1704
Qy	1114	TGCGATGAGGACCCACCATGCTGTGGCCCTGTGCGCCCTTGAGCTTCTTCAACACC	1233
Db	1705	TGGACGAAAGGCCACCATTCGTGGGGCATATCTGGCTGGCTCTTCTTCACTGTACA	1764
Qy	1234	CTGGCATTCTGGATATTCGTGAGAGCATTCACACACCCATGTTGCTGCGGCTGG	1293
Db	1765	CTGGCATTCTTTCATCTTCTGTGAGACATTTCCAGACCCATGTGCGCTCGGCGGT	1824
Qy	1234	GGCCCCATGTCTTCTGTATGCTGACACTGTCTGTGTGGCATPACATGATGATGCCGGTG	1353
Db	1835	GGCCCCATGTCTTCTGTATGCTGACACTGTCTGTGTGGCATPACATGATGATGCCGGTG	1884
Qy	1354	TACGAGGGGGCGCCAGAGTCCTCACCTGGCTCTGCGGACGAGCCCTCTTCCCTCGC	1413
Db	1885	TATGTGGGGGGCCCCACAGGTCCTTCTATGCTTGTGCGACAGGCTTCTTCAACGCTCTGC	1944
Qy	1414	TTTCAATTTTGCATCTCTCTGTATGCGCGTGTCTTCTTCCAGATGCTGTGCGCTTCAAG	1473
Db	1945	TTTCTCATCTCTCTATCTCTGATCATCACCGTGTGCGCTCTTCAATGTGTGTCTTCAAG	2004
Qy	1474	ATGGCAGAGCGCTTCCACCGGCTACAGTAACTGGGTCCGCTACCAAGGGGCCCTACTC	1533
Db	2005	ATGGCAGAGCGCTTCCACCGGCTACAGTAACTGGGTCCGCTACCAAGGGGCCCTACTC	2064
Qy	1534	TCTATGGCATTTATCACGAGTACTCAAAATGTCATTTGTGTAATGGCATGCTGTGACGG	1593
Db	2065	TTTGTGGGCTTTCATCACGGGCTCAAAAGTGTGGCCGTGTGTGTGGCAAACTGTGTGCCACC	2124
Qy	1534	CCTGATGCCACCC--CCGTACTGACCCCGATGATCCCAAGATCAATTTGTCTCTGT	1650
Db	2125	ACCATCAACCCCATTTGGCGGGACCGACCGGAGTACCCCAATATGATCTCTCTGTGCG	2184
Qy	1651	AACCCCAACTACGCAACACGCTGTGTTCACACACAGCGCTGACGTCGTGCTCATGTG	1710
Db	2185	CACCTTAACTACCGCAACGGGCTACGTTCACACACAGATGATGACTTGTCTGTCTGTG	2244
Qy	1711	GTGGGTTTCAAGCTTGTGCTTACATGTGGGCAAAAGACTGCGCCACCACTACAGAGGCCAG	1770
Db	2245	CTGGGTTTCAAGCTTGTGCTTACATGTGGGCAAAAGACTGCGCCACCACTACAGAGGCCAG	2304
Qy	1771	TTTCAACCCGCAAGACTCTTATTTTACCTCAATGCTGTCTCTCTGACACTTATG	1830
Db	2305	TTTCAACCTTCAAGACTCTTCTCTTCACTCTCTCTCTCTGACACTTATG	2364
Qy	1831	TCTGCTTACAGCGGGGTGTGTACCATTCGTGAGACCTTGTGTACTGTCTCAACCTC	1890
Db	2365	TCTGTGACAGCGGGGT	2424
Qy	1891	CTGGCCATCAAGCTGGGCTACTTGGGCCCAAGTGTCTACATGATCTCTTCTACCCGGAG	1950

[illegible]

OY	302	ACAAGCTTCACGACGACCAACGGCGGTGCTGCTGTTCTTCCGCCGACTGCATACCCTGT	361
Db	835	AGCATCTGGGCTCAGAGCCAGGACCAACCGTGTTGTGGTCTCTTAACCCGACACTGGCTGA	894
OY	362	ACCACTTCTTCAATGAGGTGCTGGCGCAAGACTTCAACGGGCGCGGTGGATCGGCTCG	421
Db	895	GAGTGTCTTCAAGTCCGTGGGTGTGGCCAACTGCATGCTGGCAAAGTGTGGTGGCTTCAG	954
OY	422	AGTCTTGGGCATGACCCCGCTCTGCACMACTCAACGAGTGGGCACTTGGCACCT	481
Db	955	AAGAAGCTGGGCATCTCCACGTACATCAACGAGCTGACTGGATCCAAAGTGTGGTGGCTTCAG	1011
OY	482	TCTTGGGCATCCATCATCAAGCGTGGCCATCCCGGGCTTCAAGTAGTTCGGCAGTGGG	541
Db	1015	TGCTCGGTGTGGCCGTCCACACGAGAACATCCCTGGGCTGAAAGGATTGAGAGTCTT	1072
OY	542	GCCCAAGGCTGGGGCGCCACCCCTCAGACGAGACCAAGCAGACTATACCTGC-----A	595
Db	1075	ATGTCAAGGCTGTACACGCTGTCTCCACGCCCTTGGCCGAGGGGTCTTGCTGACACTGA	1133
OY	596	ACCAAGAGTGGACAACTGCCCTGAACGCAACCTGTCTTCAACACCATTTCTCAGAGCTCT	655
Db	1135	ACCAAGCTGTGGCGGAGTGGCCACACGTTTCACAGCTGTATAACATGCCACGACTTGGAGCT	1192
OY	656	CTGGGGAGACGTGTGCTTAACAGCGTGTACTCTCGGCTATAGTGTGGGCCATGCCCTGC	715
Db	1195	TCTCCATGATGCGCGCTTACCAAGTGTATGAGGCTGTATACGTGTGGCCACAGGCTCTC	1255
OY	716	ACAGACTCTCTCGGCTGTGACAAAAGCAACCTGCACCAAGAGGGGTGTACCCCTGGCAGC	775
Db	1255	ACCAAGCTCTCGGAGTGTACTTCTGAGATCTGTTTCCAAGAGCCCAAGTCTACCCCTGGCAGC	1311
OY	776	TGCTTGAAGAGATGTGGAAGGTCAACTTCACTCTCTCTGGACCAACAAATCTTCTTGACC	835
Db	1315	TTCCTTCAGCAATCTTCAAAAGTGAATTTCTTCAATGAGAACTGTGGCATTTTGAAG	1372
OY	836	CGCAAGGGGAGGNTGGCTCTGACATGGAGATTTCCACAGTGGCAATGGAGCCGAGACAGA	895
Db	1375	ACAAAGGGGACACTCTAGGTTACTAGCACATCATTCGCCGTGGACTGGAAATGACCTGAAAT	1433
OY	896	ATTCCTTCCAGAGCGCTGCTCTACTACCCCCCTGACGAGCAAGCACTGAAGAAATCAACAGA	955
Db	1435	GGACCTTTGAGATCATTTGGCTTGGCCTCACTGTCTCCAGTTCATCTGGACATATAATATAGA	1492
OY	956	CATCTCT-----GCACACCGTCAACACACACATCCCTATGTCTCATGTGTTCCAAAGGT	1005
Db	1495	CAAAAATTCAGTGGCACGGGAGMACATAGGTGGCTGTGTCAAGTGTGTACACAGCACT	1552
OY	1010	GCGAGTGAAGGCAAAAGAAAGAGCGTGTGGGCAATCCACGCTGTGCTCTTCAGATGATG	1065
Db	1555	GTCCTGGCAGGGCCACCAAGAGGTGTGTGGGTTCACCACTGTCTCTTTAGATGTGTGC	1611
OY	1070	ACTGCTTCCGGGACCTTCTCTCAACCACTAGAA-----TGCCTGA	1111
Db	1615	CTCTCGAAGCTGGGAACTTTCTCAACATGATGAGCTTCAACTCTGCAGCCCTGTGGAA	1672
OY	1112	ATAACAGTGTCTTACCAAGATGAGACCTCTGTCTTCAAGGGGCAAGCTGTCTTCTG	1177
Db	1675	CAGAAGATGGGACCCCAAGAGAGACATCACTTGTCTTCCACGACAGGTGGAGTCTTGG	1733
OY	1172	AATGAGTGAAGGACCCCAACCATGCTGTGTGGCCCTGTGGCCGCGCTCGGCTCTTCAGACA	1233
Db	1735	CTTGGCATGAACCATCTCTTTGGTGTCTAATACACACTTAACAGCTATGTGTGTCTGC	1792
OY	1232	CCCTGGCCATCTGTGTGATATCTTGAAGCACTTCCAGACACCACTATGTTGCTCGGCTG	1291
Db	1795	TGTTTTGGAGCACTGCGGCTCTTGTCTGCTGCGCATTTTCAACACCTGTAGAGTCAAGCTG	1852
OY	1292	GGGGGCCCATGTGTTCTTCAATCTGTACACTGTGTGTGTGGTGTGATCAAGTGTGTCCGG	1355
Db	1855	GGGGTGAAGCTGTGCTTCTCATCTGTGGGTTCCTGTGTGGCCGGAAGTTGCACCTCTCTATA	1912

QY 1352 TGTAGCTGGGGGCGCCAGAGTCTCCACCTGCTGCGCCGACGCCCTTTCCCTCT 1411
 DB 1915 GCTTCTTGGGAGAGCCAGGTCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1974
 QY 1412 GCTTCAATTTGATCTCTGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1471
 DB 1975 GGTTCCTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2034
 QY 1472 AGATGGCCAGCCGCTTCCACGCGCTTACAGCTACTGGCTGCTGCTGCTGCTGCT 1531
 DB 2035 AGTTTCTTCAAGAGGAGCCACATTTCTACCTGCTGCTGCTGCTGCTGCTGCTGCT 2094
 QY 1532 TGTCTATGGCATTTTACAGGATCTCAAAATGCTGCTGCTGCTGCTGCTGCTGCT 1591
 DB 2095 TATTGCTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2154
 QY 1592 GGCCTTCACTCCAGCCCTTACCTGACCCGATGACCCCAAGATCAATGCTGCTGCT 1651
 DB 2155 GGCACCCAGAGCCAGCCAGGAAATACAGCGCTTCCCATGCTGCTGCTGCTGCTGCT 2214
 QY 1652 ACCCCACTACCGCAACAGCTGCTGCTTCAACACAGCCTGAGCTGCTGCTGCTGCT 1711
 DB 2215 CAGAGGTCACATCTGTAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2274
 QY 1712 TGGGTTTCACTTCCGCTTACATGGGCAAGAGTGGCCCAACATCAACAGGCGCA 1771
 DB 2275 GTACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2334
 QY 1772 TCATCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1831
 DB 2335 GGTCTCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2394
 QY 1832 CTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1891
 DB 2395 GCATTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2454
 QY 1892 TGGCCATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1951
 DB 2455 GGGGGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2514
 QY 1952 GCACACAGCCGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2007
 DB 2515 TCACATATACAGACACTTTCAGGCTTCCATCCAGACTACAGAGGCGCTGCGGC 2570
 RESULT 5
 AAD19502
 ID AAD19502 standard; cDNA; 2292 BP.
 AC AAD19502;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human novel G-protein coupled receptor (NGPCR) cDNA #2.
 XX
 KW Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
 KW antisense-therapy; signal transduction; behavioural disorder; obesity;
 KW heartbeats rate; inflammation; immune disorder; diabetes; cancer;
 KW coronary disease; ss.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT CDS 1..2292
 FT FT /*tag- a
 FT FT /product- "Human NGPCR protein"
 FT FT 85..87
 FT FT /*tag- b
 FT FT /note- "Encodes Ser"
 FT FT 880..882
 FT FT /*tag- G
 FT FT /note- "Encodes Ala"
 XX

PN W0200172842-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 28-MAR-2001; 2001WO-US09996.
 XX
 PR 28-MAR-2000; 2000US-192978P.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abulin A;
 DR Zambrowicz B, Sands AT;
 DR WPI: 2001-616474/71.
 XX
 DR P-PSDB: AAE11970.
 XX
 PT Novel isolated polynucleotides encoding human G protein coupled
 PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
 PS expressed NGPCRs for diagnosis of disease, and as probes or primers -
 XX
 PS Claim 4; Page 74-75; 80pp; English.
 CC
 CC The present sequence is a cDNA encoding human novel G-protein coupled
 CC receptor (NGPCR) protein. NGPCRs are transmembrane proteins that span
 CC the cellular membrane and are involved in signal transduction after
 CC ligand binding. The NGPCR polynucleotide sequences, are useful
 CC in diagnosis and treatment of a disease involving NGPCR, for detecting
 CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NGPCR in
 CC the body or abnormalities in the signal transduction pathway mediated by
 CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
 CC trial monitoring and/or the treatment of physiological (heartbeats rate)
 CC or behavioural disorders. NGPCR is useful for identifying compounds
 CC useful in the therapeutic treatment of obesity, inflammation, immune
 CC disorders, diabetes, heart and coronary disease, metabolic disorders and
 CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
 CC genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests.
 SQ Sequence 2292 BP; 437 A; 690 C; 638 G; 524 T; 3 other;
 Query Match 19.2%; Score 386.4; DB 22; Length 2292;
 Best Local Similarity 51.1%; Pred. No. 1.5e-71;
 Matches 1042; Conservative 1; Mismatches 947; Indels 48; Gaps 4;
 QY 1 ATCACCCTACAGCGCCATCAGCGATGAGTCGAGACAGAGTGGCTTCCGCTTGGCTG 60
 DB 265 ATTAGCTATGCGGCGCCAGGAGCGAGCTGAGGTGAAGGCGAGTATCCCTTTCCG 324
 QY 61 CGTACCACACCCAGCGCCGACACACAGCTGAGAGCCATGCTGATGCTGCACTTC 120
 DB 325 CGCACCATCCCATATACAGTACAGTACGAGTACAGTACGAGTACGAGTACGAGTAC 384
 QY 121 CGCTGAACTGATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 DB 385 GGTGTGACTGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 444
 QY 181 CTGCTTGGGAGGCGCTGCGCGCGCGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCT 240
 DB 445 GCACTGAGACAGAGCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 504
 QY 241 ACACTGACACCCACAGACATGACGTGAGAGAGAGCCGAGCGCTGCTGCTGCTGCTG 300
 DB 505 -----TTCTCTGCGCAGAGTGGCGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 546
 QY 301 GACAAGCTGACAG 360
 DB 547 CGCAGCTTGGCCAGAGCGCGGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606
 QY 361 TACCACTTCTGATGAGTGGTGGCGCAGAACTTACAGGAGGCGCGCTGCTGCTGCTGCT 420

Db 607 AGGTTGTTTTGAGTCGCGTGGTGTGACCAACTGACTGGCAGAGTGGGTCGCTCA 666
 QY 421 GAGTCTGGGCGCATGACCGGCTGCTGCACAACCTGACGAGAGCTGGGCACTTGGGCACC 480
 Db 667 GAGGCTGGGCGCTCTCCAGGACATCACTAGGGGTGGCCGGGATCCAGCGCATTTGGGATG 726
 QY 481 TTCTGGGCAATCAACATCCAGAGCGTCCCATCCGCGCTTCAGTAGAGTTCCGCGAGTGG 540
 Db 727 GTGCTGGGCGTGGCATCCAGAGAGAGGCTGCTCGGCTGAGAGCGCTTGAAGAGAGCC 786
 QY 541 GGGCCACAGGCTGGGCGCCGACACCTCCAGCAGAGACAGCAGAGTAT-----ACCTGC 594
 Db 787 TATGCCCCGGGACAGACAGAGAGGCCCCCTAGGCTTCACAGAGGCTCCTGTGACAGAGC 846
 QY 595 AACCGAGAGTGGACAACTGCTGAACGCCACCTGTCTTCACAACTTCAGAGCTC 654
 Db 847 AATCAGCTCTGAGAGAGATGCCAACCTTTCATGRCACACAGATCCCAAGTCAAAAGCC 906
 QY 655 TCTGGGAGAGCTGTCTGCTACAGCGTGTACTCTGCGGTCTATGCTGTGGCCGAGCCCTG 714
 Db 907 TTTCTCATGAGTTTCTGCTACACGCAATACGGGCTGTGTATGCGGTGAGCCATGGCCTC 966
 QY 715 CACAGCTCTCGGCTGTGACAAAAGCACTGACCAAGAGGAGTGTACCCCTGGGAG 774
 Db 967 CACGAGCTCTGGGCTGTGCTGCTGAGCTGTGTTCCAGGGGCGAGTCTACCCCTGGGAG 1026
 QY 775 CTGCTTGAAGAGATCTGAGAGGTCACTTCACTCTCTGACACCAAACTTCTTCGAC 834
 Db 1027 CTTTGGAGAGATCCACAGAGTGCATTTCTCTACAAAGAGACAGTGGCGTTTAT 1086
 QY 835 CCGCAGAGGAGTGGCTGTGCACTTGGAGATTTGCCAGTGGCAATGGAGCCGAGCCAG 894
 Db 1087 GACCAACAGAGATCCCTCAGTATGATTAACATTAATGGCTGGAGCTGAGATGAGCCAG 1146
 QY 895 AATCCCTTCCAGAGGCTGGGCTCTAC-----TACCCCTGACAGAGAGCTGGAAGAC 948
 Db 1147 TGGACCTTCCAGGCTGTGCTTCCACAGAGTGTCTCCAGTTCCTAAACATTAATGAG 1206
 QY 949 ATCAAGACATCTCTGACACAGCTCAACAACAGATCCCTATGTCATGTGTTCCAAGAG 1008
 Db 1207 ACCAAATTCAGTGGCAGGAGAGAGACAAACAGTGGCTAGTGTGTGTTCCAGCCAG 1266
 QY 1009 TGGCAGTCAGGGCAAAAGAGAGAGCGTGTGGGATCCAGCTGTGCTTCGATGTCATC 1068
 Db 1267 TGTCTTGAAGGGCAGCAGCGAGTGTGAGGGTTTCCATCACTGCTTGAAGTGTG 1326
 QY 1069 GACTGCTTCCCGGACCTTCTCAACCACTGAA-----TGCCCG 1110
 Db 1327 CCTTGTGGGGCTGGGACCTTCTCAACAAGAGTACTCTTCAGATGCCAGCTTGTGGG 1386
 QY 1111 AATTAAGAGTGTCTTACCAAGAGTGAACCTTCTCTCAAGCGGAGCTGTGTTCTCTG 1170
 Db 1387 AAAGAAAGTGGGACGTCGAGGAAGCCAGACCTTCCGCGGCAAGTGTGTTTGG 1446
 QY 1171 GAATGGCATGAGGACCCCACTGCTGTGGCCCTGTGCGCCCTGCGCTTCTCTACAC 1230
 Db 1447 GCTTGTGAGGACACACTTGTGGTGTGCTGAGAGTAAACCTCTGCTGCTGTG 1506
 QY 1231 ACCCTGGGCATTCGATATATTCGAGAGGACCTTCAGACACCATAGTTGCGTGGGT 1290
 Db 1507 CTGCTTGGGAGCTGCTGGCTGTGTTGCTGAGACCTAGACACCCCTGTGTGAGGTACGA 1566
 QY 1291 GGGGGCCCCATGCTGCTCTGATGCTGACACTGCTGTGAGGAGCAATAGTGTGGTCCCG 1350
 Db 1567 GGGGGCGCGCTGTCTTCTTATGCTGGGCTCCGAGGAGAGTGGCAGCTCTAT 1626
 QY 1351 GTGACGTGGGCGCCCAAGGTCTCACTGCTGTGCGCGCAGAGCCCTCTTTCCCTC 1410
 Db 1627 GGGCTCTTTGGGAGAACCAAGGCTGTGCTGTGCTGACGCAAGCCCTCTTCCCTT 1686
 QY 1411 TGCTTCACAATTTGATCTCTGATGCGCGGTCTTTTCCANATGCTGTGGCCCTTC 1470
 Db 1687 GGTTCACCACTTCTCTGCTGCTGAGAGTGGCTCATTTCCAACTAATCATCATCTTC 1746

QY 1471 AAGATGGCCAGCCGCTTCCACAGGCGCTACAGCTACTGCTGCTCCGCTACAGAGGCGCTAC 1530
 Db 1747 AAGTTTCCACCAAGTACTACATTTTACACAGCGCTGGGTCCAAACACAGGCTGTGCTGC 1806
 QY 1531 GTCTCTATGACATTTATCAGGTAATCAAAATGCTATGTTGTTAATTTGATGCTGCA 1590
 Db 1807 CTGTTTGTGAGATGACGTACAGCGGCCAGCTGCTTATCTGTCAACTTGTGCTGGGTG 1866
 QY 1591 CGGCTCAGTCCACCCCGCTACAGCCGATGACCCCAAGATCACATTTGTCTCTGT 1650
 Db 1867 TGGACCCCACTGCTGTGAGGAATACAGCGCTTCCCACTGTGTGATCTGTGAGTGC 1926
 QY 1651 AACCCCACTACCCGACAGCGCTGCTTTCAACACACAGCTGAGACCTGCTGCTCAAGTG 1710
 Db 1927 ACAGAGCACTACCTCCGCTTCTACTGAGCTTCTCTCAATAGGCTCTCTCTCCATC 1986
 QY 1711 GTGGTTTACGCTTGCCTACATGGGCAAGAGCTGCCCCAACACTACAGAGGCCAAG 1770
 Db 1987 AGTGGCTTGGCTGACGCTACCTGGGTAGAGACTTGGCCAGAGAACTACAGAGGCCAAA 2046
 QY 1771 TTTATCAACCTCAGCATGACCTTCTATTTCACTTCACTGCTGCTGCTGACCTTCATG 1830
 Db 2047 TGTGCTACCTTCAAGCTCTCTCTTCACTTCTGCTGTGATGCTTCTTCAACAGGCTC 2106
 QY 1831 TCTGCTACAGCGGGGCTGCTGATCAACATGCTGACCTTGTGCTGCTCAACCTC 1890
 Db 2107 AGGCTACAGAGGCAAGTACCTGCTGCGGCAACATGATGGCTGGGCTGAGGAGCTG 2166
 QY 1891 CTGGCCATCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1950
 Db 2167 AGCAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2226
 QY 1951 CGCAACAGCGCGCTTCTTCAACAGATGATCAAGGCTGACACATGAGAGGAGT 2008
 Db 2227 CTCAACAGCAGAGACTTCCAGGCTTCACTTCAAGACTACAGAGGCGCTGGGCT 2284

RESULT 6
 AAD19501
 ID AAD19501 standard; cDNA: 2526 BP.
 AC AAD19501;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human novel G-protein coupled receptor (NGPCR) cDNA #1.
 XX
 KW Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
 KW antisense-therapy; signal transduction; behavioural disorder; obesity;
 KW heartbeat rate; inflammation; immune disorder; diabetes; cancer;
 KW coronary disease; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..2526
 FT CDS
 FT unsure
 FT /product- "Human NGPCR protein"
 FT /tag- a
 FT /tag- b
 FT /note- "Encodes Ser"
 FT 319..321
 FT /tag- c
 FT /note- "Encodes Ala"
 XX
 PN W0200172842-A2.
 XX
 PD 04-OCT-2001.
 XX
 PR 28-MAR-2001; 2001WO-US09996.
 XX
 PR 28-MAR-2000; 2000US-192978P.

XX
PA (LEXI-) LEXICON GENETICS INC.
XX

PI Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abuin A,
PI Zambrowicz B, Sands AT;

DR WPI; 2001-616474/771.

DR P-PSDB; AAE11969.

PR Novel isolated polynucleotides encoding human G protein coupled
 PR receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
 PT expressed NGPCRs for diagnosis of disease, and as probes or primers -
 XX
 PS Claim 1, Page 71-72, 80pp; English.
 PS

CC The prent sequence is a cDNA encoding human novel G-protein coupled
CC receptor (NGPCR). protein. NGPCRs are transmembrane proteins that span
CC the cellular membrane and are involved in signal transduction after
CC ligand binding. The NGPCR polynucleotide sequences, are useful
CC in diagnosis and treatment of a disease involving NGPCR, for detecting
CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
CC disease, for screening drugs effective in treatment of symptomatic or
CC phenotypic manifestations of perturbing the normal function of NGPCR in
CC the body or abnormalities in the signal transduction pathway mediated by
CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
CC trial monitoring and/or the treatment of physiological (heartbeats rate)
CC or behavioural disorders. NGPCR is useful for identifying compounds
CC useful in the therapeutic treatment of obesity, inflammatory compounds
CC disorders, diabetes, heart and coronary disease, metabolic disorders and
CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
CC genomic library which is helpful for identifying polymorphisms,
CC determining the genomic structure of a given locus/allele and designing
CC diagnostic tests.

SD Sequence. 2526 BP; 469 A; 770 C; 697 G; 587 T; 3 other;

Query Match	19.2%	Score 386.4	DB 22	Length 2526
Best Local Similarity	51.1%	Pred. No. 1.6e-71		
Matches 1042; Conservative	1	Mismatches 947	Indels 48	Gaps 4

QY	1	ATCCGACAGAGCGGCATCAGAGATGAGACTGCGAGACAAAGGTGGCGCTCCGGGCTTGGCTG	60
Db	499	ATTAGCTATGCGGCCACAGACAGGAGACCTCAGCTGAAAGGGCAGTATCCCTCTTTCCTG	558
QY	61	CGTACCAACCCACAGCGCCGACCAACACAGTCGAGGCCATGGCTGCAGCTCATGCTGCATTC	120
Db	559	CGCACCATCCCCATGACATGACAGTACCAAGGTGAGACCATGGTGGCTGCTGCTGCAGAAAGTTC	61.8
QY	121	CGCGGGAACCTGAGATCATTTGGCTGTGTGATGAGACGACACACTATATGGCCGCGCAATGGCCAG	180
Db	619	GGGGTGGACCTGGAGTCTCTCTGGTTGGCAGCAGTACACATATGGGACGCTAAGGGGTGCAG	67.8
QY	181	CTGCTTGGCAGGCGCTGTGGCCCGGCGGACATCTGCATCCCTTCCAGAGACGCTGCC	240
Db	679	GCACTGGAGAACCAGGCGCATGGTGCAGGGGATCTGCATTTCCTTTCAAGAGATCATATCCC	73.8
QY	241	ACACTGACGCCCAACGAAACATGACTGCACAGAGAGCCGCAAGCGCCTGTGTACCAATTGTG	300
Db	739	-----TTCTCTGCCAAGTGGCGCATGAGAGATGACGTACCTCATG	780
QY	301	GACAGCTCGACAGACAGCAGCGCGGCTGTGGTCTGTTCGCGCCGACCTGCACCTG	360
Db	781	CGCCACTTGGGCCAGGCGCGGGGCCACCGTGTGTGTTTTCATCAGCCGGCATTTGGCC	840
QY	361	TACCACTTCTCAATGAGGTGCTGCGCAAACTTCAAGGCGCGCGCTGTGGATGCGCTCC	420
Db	841	AGGGTGTTCGATGCGTCCGTGTGCTACCAACCTGTCTGGCAAGGTGTGGGTGCGCTCA	900
QY	421	GAGTCTCGGGCAATCGACCGGTCCTGCACAACTCAGGAGGTCGGGCCACTTGGGCACC	480
Db	901	GAACTCTGGGGCTCTCTCAGGACATATCACTGGGGTGGCCCGGGATCCAGCCCATTTGGGATG	960
QY	481	TTCTCGGACATCAACATCCAGAGCGTGCACATCCCGGGCTTCAGTGAATTCGCGAGTGG	540

Dp	961	GTGCTGGGGCGTGGCCATTCAGAGAGAGGGCGTGCCTCGGCGCTGAAGGCGCTTGTGAAGACC	1020
Oy	541	GGCCCAAGAGCTGGGGCCGACCCCTTCAGCAGGACCAAGCAAGCTAT-----ACCTGC	594
Dp	1021	TATGCCCCGGGAGCAAGAGAGGCCCTTAGGCCCTTGGCAACAAAGGCTCTGTGCGCAGC	1080
Oy	595	AACCAAGAGTGGCGCAACTACCTCTGAACAGCACCCTTGTCTTAAACAACTTCTCAGGCTC	654
Dp	1081	AATCAGCTCTGCAGAGATATCCCAAGCTTTCATSRCAACACAGATGCCAAGCTCAAAACC	1140
Oy	655	TCCTGGGAGAGCTGTCTCTACAGCGTGTACTCTGGGGTATATGTGTGGCCAGCCCTG	714
Dp	1141	TTTCATCATGATTTCTGCTCTCAACAGCATACCGGGCTGTGTATGCGGTGGCCATGGCTC	1200
Oy	715	CACAGCCCTCCGCGCTGTGCACAAAGACCTGTGACCAAGAGGGTGTATACCCCTGGGAG	774
Dp	1201	CACCAAGCTCTCGGGCTGTGCTCTGTGAAGCTTGTTCAGGGGGCCGAGTCTACCCCTGGAG	1260
Oy	775	CTGCTGAGAGATCTGGAAAGTCAACTTCACTCTCTGTGACACCAAAATCTTCTTCAC	834
Dp	1261	CTTTTGGAGCAGATCCAGAAAGTGATTTTCCTCTACACAAAGACACTGTGGCGTTTAAT	1320
Oy	835	CCCGAAGGAGACGTGGCTGTGCACCTTGGAGATTGTCAATGGCATGGGACCGAGCGAG	894
Dp	1321	GACAACAGAGATCCCTCACTAGTACTATPACATPATTTGCTCGGAGATGAAATGGACCCAG	1380
Oy	895	AATCCCTTCAGAGCGTGGCTGCTCTAC-----TACCCCTCAGCAGCAGCTGAAGAC	948
Dp	1381	TGAGACCTTCACAGGTCCTGGTTCTCTCACATGCTCCAGTTCACCTAACTAAATAG	1440
Oy	949	ATCAAGACATCTTGCACACCGTCACAAACACAGATCCGTATGTCCATGTGTCCAGAGG	1008
Dp	1441	ACCAAAATCCAGTGGCGACGGAAAGAACAAACAGTGTCTCACTGTGTGTCCAGCGAC	1500
Oy	1009	TGCAGTCAAGGCAAAAGAGAGACCTGTGGGACTCAGCTGTGCTGTGAGAGTCAATC	1068
Dp	1501	TGCTTTGAAGGCAACAGGAGTGTGTACGGGTTTCCATCATGCTGTTGAAGTGTG	1560
Oy	1069	GACTGCTTCCCGGCACTTCTCACAACCACTGA-----TGCCG	1110
Dp	1561	CCCTGTGGGGCTGTGGACCTTCTCAACAAGAGTGAACCTTACAGATGCCAGCCTGTGGG	1620
Oy	1111	AATPACGATGGTCTPACAGAGTGAAGCTCTGTGTTAAACGGGACAGTGGTCTCTCG	1170
Dp	1621	AAABAAGATGGACACTGAGGAGAACCAAGACTGTCTCCGCGACATGTGTGTTTTG	1680
Oy	1171	GAATGCGATGGGACCCACCACTGCTGTGGCCCTGTGTCGGCCCTGGGCTTCTCAGC	1230
Dp	1681	GCTTTGGGTGAGCAACCTGTGGGATGTGTGGAGCTAACAAGCCTGTGTGTGTGTG	1740
Oy	1231	ACCTGGGCATCTCGTGTGATATTTCTGAAGGCACTTCCAGACACCATAAGTGTGCTGCT	1290
Dp	1741	CTGTGTGGGACTGTGTGCCCTGTGTGCTGTGACCTGACACCCCTGTGTGTGAGTCAGCA	1800
Oy	1291	GGGGAGCCCAAGTGCCTTCAGATGTGACACAGTGCCTGTGCTGGCAATAGTGGTCCCG	1350
Dp	1801	GGGGGCGGCTGTGTCTTTTAATGTGGGCTCCCTGTGACAGAGTATGGGACCTCTAT	1860
Oy	1351	GTGTAGCTGGGGCGCCCAAGGCTCTCACCTGTGCTTGGCCGAGGCGCTTTCCTCCTC	1410
Dp	1861	GGCTTCTTTGGGGAACCAACAAGGCGCTGGGTGCTGTACGCGAGGCGCTTTCCTTT	1920
Oy	1411	TGCTTCACATTTGCATCTCTCTATATGCGCGTGTCTTTTCCAAATGTGTGCGGCTTC	1470
Dp	1921	GGTTTACACATCTTCTGTGTCTCTCTGAGAGTGTGGTCAATTCATATCATATCTTC	1980
Oy	1471	AAGATGGCCAGCGCTCCACAGCGCTCAAGCTACTATGGGTCCGGTACAGAGGGCGCTTAC	1530
Dp	1981	AAGTTTTCACCAAGTACTATCATTTCTACACAGCTGGGTCTCAAAACCAAGCGTGTGCG	2040
Oy	1531	GTCTTATGGCAATTATCAGGTAAGTCAAAATGTGTCATTTGTGTGAATTTGGATGTGGCA	1590

Db 2041 CTGTTTGATGATCAGCTCAGCGGCCAGCTGCTTATCTGTCTAATCTGGCTGTGTG 2100
 QY 1591 CGGCTTACGTCACACCCCGTACTGACCCCGATGACCCCAAGATCAATTTGTCCTGT 1650
 Db 2101 TGGACCCACATGCTGCTGATGAGGAATACAGGCTTCCCGCATCTGTATGCTTAGTGC 2160
 QY 1651 AACCCCACTACCCGACAGCCCTGCTGTTCAACACAGCCTGAGCCTGCTGCTCACTG 1710
 Db 2161 ACAGAGCAACTCCCTGGGCTTCACTGAGCTTCTCTCAATAGGCTCTCTCAATC 2220
 QY 1711 GTGGTTTACGTTGCTTACATGAGGCAAGAGCTGCCCACTACAGAGGCCAAG 1770
 Db 2221 AGTGGCTTGGCTGACATCTGAGTGAAGCACTTCCAGAGAACTCAACAGAGGCCAAA 2280
 QY 1771 TTCACTACCCCTGAGTACCTTCTATTTCACTTCACTTCCCTGCTGACCTTCAATG 1830
 Db 2281 TGTGCTACCTTACGCTCTGCTTCACTTCTGCTGATGCTTCTTCAACAGGCC 2340
 QY 1831 TCTGCTTACAGCGGGGCTGCTGACATGCTGACCTTGTGCTGCTCACTGCTC 1890
 Db 2341 AGCGTCTACAGAGCAATGCTGCTGCGGCAACATGATGCTGAGCTGAGCAGCTG 2400
 QY 1891 CTGGCATCAGCTGCTGCTTCTGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTG 1950
 Db 2401 AGCAGCGCTTGGTGGTATTTCTGCTAGTCTAGTCTAGTCTGCTGCTGCTGCTG 2460
 QY 1951 CGCAACAGCGCCGCTTCTTCAACAGATGATCCAGGCTACACCATGAGAGGAGCT 2008
 Db 2461 CTCACAGCAGACAGACACTTCCAGGCTCTCANTGAGCTACAGAGGCGCTGCGGCT 2518

RESULT 7
 AAD19505
 ID AAD19505 standard; DNA; 2951 BP.

AC AAD19505;
 DT 18-DEC-2001 (first entry)
 DE Human novel G-protein coupled receptor DNA with 5' and 3' regions.
 KW Human; G-protein coupled receptor; GPCR; gene therapy; drug screening; antilease therapy; signal transduction; behavioural disorder; obesity; heartbeate rate; inflammation; immune disorder; diabetes; cancer; coronary disease; ds.
 OS Homo sapiens.
 PN WO200172842-A2.
 PD 04-OCT-2001.
 PE 28-MAR-2001; 2001WO-US09996.
 PR 28-MAR-2000; 2000US-192978P.
 PA (LEXI-) LEXICON GENETICS INC.
 PI Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abulin A;
 PI Zambrowicz B, Sands AT;
 DR WPI; 2001-616474/71.
 PT Novel isolated polynucleotides encoding human G protein coupled receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately expressed NGPCRs for diagnosis of disease, and as probes or primers -
 PS Disclosure; Page 79-80; 80pp; English.
 CC The present sequence is human novel G-protein coupled receptor (NGPCR) DNA with 5' and 3' regions. NGPCRs are transmembrane proteins that span the cellular membrane and are involved in signal transduction after ligand binding. The NGPCR polynucleotide sequences, are useful

CC in diagnosis and treatment of a disease involving NGPCR, for detecting
 CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NGPCR in
 CC the body or abnormalities in the signal transduction pathway mediated by
 CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
 CC trial monitoring and/or the treatment of physiological (heartbeate rate)
 CC or behavioural disorders. NGPCR is useful for identifying compounds
 CC useful in the therapeutic treatment of obesity, inflammation, immune
 CC disorders, diabetes, heart and coronary disease, metabolic disorders and
 CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
 CC genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests.

SO Sequence 2951 BP; 559 A; 888 C; 831 G; 669 T; 4 other;

Query Match 19.28; Score 386.4; DB 22; Length 2951;

Best Local Similarity 51.18; Pred. No. 1.6e-71; Matches 1042; Conservative 1; Mismatches 947; Indels 48; Gaps 4;

QY 1 ATCACTTACAGCGGCTGATGAGTGGAGAGCAAGTGGCTTCCGCTTGTGCTG 60
 Db 730 ATTAGCTATGCGGCGACGACGAGACGCTCAGCTGAAGCGGCAATCTCTTTCTG 789
 QY 61 CGTACACACCCAGCGCCGACACACGCTGAGGCAATGTCAGCTGATCTGCACTTC 120
 Db 790 CGCACATCCCAATGACATGACATACAGGTGAGACATGATGCTGCTGCTGCAAGATTC 849
 QY 121 CGCTGGAATGATCTGCTGCTGAGCAGCAGCACTTATGCGCGGCAATAGCCAG 180
 Db 850 GGGTGGACCTGATCTCTCTGTTGGCAGAGTACGATGAGGCGAGCTAAGGGGTGAG 909
 QY 181 CTGCTTGGCAGCGGCTGCGGCGGCGACATGTCATGCTGCTTCCAGAGACCTGCCC 240
 Db 910 GCATGAGGAACAGGCGCAGCTGTCAGGGGATCTGCAATCTTTCAGAGATATATGCC 969
 QY 241 ACACTGACGCCAACACAGACATGACGTAGAGAGCGCCAGCGCTGATGACATTTGTG 300
 Db 970 -----TTCTGCGCCAGGTGGGCGATGAGAGATGCACTGCTCATG 1011
 QY 301 GACAAGCTGACAGCAGAGACAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 Db 1012 CGCCACCTGCG 1071
 QY 361 TACCACTTCTCAATGAGTGTGCTGCGCAGAACTTCAAGCGCGCGCGCGCGCGCGCG 420
 Db 1072 AGGGTGTGTTTTCAGATCCGTGTGCTGACCAACTGACTGACAGAGTGTGCGCTCA 1131
 QY 421 GAGTCTGGGCGCATGACCGCGGTCTGACACACTGACGAGAGCTGGGCGCACTTGG 480
 Db 1132 GAGGCTGGGCGCTCTCCAGCAGCATCACTGGGTGCGCGGAGTCCAGCGCATTTGG 1191
 QY 481 TTCTGGGCGATCAGCATCAGAGCGTGGCCATCCCGGCTTCAATGATGCTCCGAGAG 540
 Db 1192 GTGCTGGGCGTGGCGATCAGAGAGGCGTGTCTCTGCTGAGAGCGCTTGAAGAGCC 1251
 QY 541 GGGCCACAGGCTGGGCGCGCCACCTTCAAGAGAGCAGAGCGAGACTAT-:-----ACCTGC 594
 Db 1252 TATGCGCGGGGAGACAGAGAGGCGCCCTTAGCGCTTGCACAGAGGCTCTCTGTGAG 1311
 QY 595 AACCAAGAGTCCGACACTGCTGAGAGCGCACTTGTCTTCAACACCATTTCTAGGCTC 654
 Db 1312 AATAGCTCTCTCAAGATGCAAGCTTCTCATGACACACAGAGAGCCCAACCTCAAGCC 1371
 QY 655 TCTGGGAGCGTGTGCTTACAGGCTGACCTGCGGTCTTGTGCGGCGGCGGCGGCTG 714
 Db 1372 TTCTCATGATTTCTGCTTACAGCATACCGGCTGTGATGAGGCGCCATAGGCTCT 1431
 QY 715 CACAAGCTCTGCGCTGTGACAAAAGCACTGCAACAGAGGCTGTCTACCCCTGGCAG 774
 Db 1432 CACCAAGCTCTGCGCTGTGAGAGCTTGTTCAGAGGCGGCGAGTCTACCCCTGGCAG 1491

QY	775	CTGCTGAGGAGATCTGGAGAGTCACTTCACGCTCTGGACACCAAACTCTCTTGAC	834
Db	1492	CTTTGGAGAGAGATCCAGCAAGGTGCTTTCCTTCTACACAGAGAACTGTGGCTTTAT	1551
QY	835	CCGCAAGGGAGCGTGCTCTGCACCTTGGAGATTGTCCAGTGGCAATGGAGCCGAGCCAG	894
Db	1552	GACAAACAGAGATCCCTCACTAGCTATTAACATATATATGCTGGAGCTGGAAATGGACCCAG	1611
QY	895	AATCCCTTCCAGAGCGTGCCTCTAC-----TACCCCTCTACAGCAGCTGAAGAC	948
Db	1612	TGAGACTTCAAGGCTCCGAGGTCTCTCCACATGGTCTCCAGTTTCACTAAACATAATATAG	1671
QY	949	ATCAAGACATCTCTGCACACCGTCAACAAACAGATCCCTATGTCCATGTGTCCAAAGG	1008
Db	1672	ACCAAAATCCAGTGGCACAGGAAAGGACMACAGATGCTTAATCTGTGTCTCCAGCAC	1731
QY	1009	TGCCAGTACGGGCAAAAGAAAGAAAGCTGTGGGCAATCCAGTGTGCTTGGATGATC	1068
Db	1732	TGCTTTAAGGGACCCAGCAGAGTGGTTAAGGGGTTTTCATCACTGCTGTTGAGATGTG	1793
QY	1069	GACTGCTTCCCGGACACTCTCCACACCACTGAA-----TGCCG	1110
Db	1792	CCCTGTGGGCTGGGACCTTCTCTCAACAAAGATGACCTTACAGATGCCAGCTTGTGG	1851
QY	1111	AATACAGATGTGTCTTACACAGATGAGACCTCTGCTTCAAGCGGACGTGTCTTCTG	1170
Db	1852	AAACAAAGTGGGCACTGTAGGGAAACCAAGACCTGCTTCCCGCACTGTGTGTTTGG	1911
QY	1171	GAATGGATGGGCAACCCACATCGGTGGGGCCGTGCTGGGCGCCCTGGGCTTCCAGC	1230
Db	1912	GCTTGTGGGTGACACACCTTGTGGGTGTGCTGTGGACGTAAACACGTGTGTGCTGTG	1971
QY	1231	ACCTGGCCATCTGTGTATATCTGAGAGCACTTCCAGACACCCATATGTCGTGCT	1290
Db	1972	CTGCTTGGGACTGCTGGCTGTTGCTGTGGACCTAGACACCCCTGTGGTGAAGCACA	2031
QY	1291	GGGGGGCCCATGTGTCTTGATGTGACATGTCTGTGTGGCATACATGTGTGCCG	1350
Db	2032	GGGGGGCCGCTGTCTTATGTGTGGGCTCCCTGGACGAGTATGGACGCTCTAT	2091
QY	1351	GTGTACGTGGGGCGGCCAAGGTCTCACACTGGCTGTGGCGGAGGCCCTCTTCCCTC	1410
Db	2092	GGCTTCTTTGGGGAACCCACAAGGCGTGGGTGTCTACGGCAGGCCCTCTTGCCCTT	2151
QY	1411	TGCTTCACAATTGTGCATCTCTGTATGCGCTGTCTTCTTCCAGATGCTGCGCTTC	1470
Db	2152	GCTTTCACCACTTCTGCTCTCTCTCGACAGTTCGCTCATTCACACTATATCATCTTC	2211
QY	1471	AAGATGGCCAGCGCTTCCACGCGGCTACAGCTACTGGGTGCGCTACCAAGGGGCCCTAC	1530
Db	2212	AAGTTTTCACCAAGTACTTACATCTTACACAGCGCTGGGTCCAAACCAAGGTGTGCG	2271
QY	1531	GTCTATGGCAATTATCAAGTACTCAAAATGTCTATGTGTATTTGGCATGTGTGCA	1590
Db	2272	CTGTTTGTATGATACACTCAGGGGCCACAGCTGTATCTGTCTAATCTGGCTGTGTG	2331
QY	1591	CGGCGTCAAGTCCACCCCGTACTGTACCCCGATGACCCCAAAATCAATATGTCTCTGT	1650
Db	2332	TGGAGCCCACTGCTGTGTAGGAATACAGCGGCTTCCCACTGTGTATGTGTGAATGTC	2391
QY	1651	AACCCCACTACGCAACACCTGTCTGTTCACACACAGCGCTGGACTGTGCTCACTG	1710
Db	2392	ACAAGACAACTCCGTGGGCTTCAATCTGGGCTTCTCTACAAATGGGCTCTCTCATTC	2451
QY	1711	GTGGGTTTCAAGCTTTCGCTTACATGGGCAAAAGAGTGTCCACCACTACAGAGGCGAAG	1770
Db	2452	AGTGCCTTGGCTGCACTACCTGTGGGTAAAGGATGTCCAGAACACTACAGAGGCGCAA	2511
QY	1771	TTTCAACCCGTCAGATGACCTTCTATTTTCACTATATCCGCTCTGCAACCTTATG	1830
Db	2512	TGTGTACACTTACAGCTGCTTCTTCAACTGTGTGTCTGTATGTGCTTCTTCAACAGGCC	2571
QY	1831	TCTGCTTACAGCGGGGTGTGTACCATCTGTGGACCTTGTGGTCACTGTGTCAACTC	1890

Db	2572	AGCGTCCTACGACGGCAGTACCTGCGGCGCCAAATATATGCTGGCTGAGCAGCTG	2631
Qy	1891	CTGGCCATCATCCCTCGGCTACTTCGCGCCCAAGTGTACATGATCCTCTTCTACCCGGAG	1950
Db	2632	AGCAGCGCGCTTCGGTGGGATATTTCTCCCTAAATGTCTACGTATCTCTGCGCCAGAC	2691
Qy	1951	CGCAACACGCGCCGCTACTTCACAGATGATCTCAGGCTACACCATATGAGGAGACT	2008
Db	2692	CTCAACACGACACGAGGACTCTCAGGCGCTCCATTCACGAGACTACACGAGGCGTGGCT	2749
RESULT 8			
ABL55953			
XX	ABL55953	standard; cDNA; 2759 BP.	
XX	ABL55953;		
XX	17-JUN-2002	(first entry)	
DE	Human G-protein coupled receptor encoding cDNA SEQ ID NO 13.		
XX	Human: GPCR, G-protein coupled receptor; receptor; anti-HIV; antitumour		
KW	antihistaminergic; antiallergic; antianemic; antiaesthetic; vitruide;		
KW	immunosuppressive; dermatological; nephrotoxic; antisynt; antithyroid;		
KW	cytotoxic; neuroprotective; osteopathic; antipruritic; antineumatic;		
KW	antiarthritic; thymomelic; antitumor; ophthalmologic; antibacterial;		
KW	fungicide; antiparasitic; protozoicide; antihelminthic; antidiabetic;		
KW	antilepterosclerotic; hepatotropic; anticonvulsant; anorectic; metabolic;		
KW	antiemetic; antidrinoetic; neuroleptic; cerebroprotective; nootropic;		
KW	antiparkinsonian; depilatory; tranquilizer; hypotensive; vasotropic;		
KW	cardiac; antidiagonal; vulnery; proliferative disorder; cancer;		
KW	neurological disorder; Alzheimer's disease; Huntington's disease;		
KW	Parkinson's disease; multiple sclerosis; meningitis; prion;		
KW	cardiovascular disorder; acquired immunodeficiency syndrome; AIDS;		
KW	Crohn's disease; diabetes mellitus; rheumatoid arthritis; gene; ss.		
XX	Homo sapiens.		
OS			
XX			
FH	Key	Location/Qualifiers	
FT	CDS	53..2578	
FT		/*tag= a	
FT		/product= "G-protein coupled receptor"	
XX			
XX	WO200198323-A2.		
PN			
XX			
PD	27-DEC-2001.		
XX			
PF	15-JUN-2001; 2001WO-US19354.		
XX			
PR	16-JUN-2000; 2000US-212483P.		
PR	23-JUN-2000; 2000US-213950P.		
PR	26-JUN-2000; 2000US-214062P.		
PR	07-JUL-2000; 2000US-216595P.		
PR	14-JUL-2000; 2000US-218936P.		
PR	19-JUL-2000; 2000US-219154P.		
XX			
PA	(INCY-) INCYTE GENOMICS INC.		
XX			
XX	Ial P, Graul R, Hafalia AJN, Walla NK, Thornton M, Nguyen DB;		
PI	Li Y, Gandhi AR, Patterson C, Kalllick DA, Baughn M, Ramkumar J,		
PI	Tribouley CM, Lee EA, Ding L, Burford N, Yao MG, Yang J,		
PI	Griffin JA;		
XX			
DR	WPI: 2002-139780/18.		
XX	P-PSDB: ABB77319.		
XX			
PT	Novel G-protein coupled receptor protein and polynucleotides useful for		
PT	diagnosing, treating or preventing disorders of cell proliferation e.g.		
XX	cancer, neurological and genetic disorder e.g. thalassemia		
XX			
PS	Claim 50; Page 117-118; 121p; English.		


```

Db 1913 GGCCTTTGGGGAACCAAGGCGTGTGCTTGTACGCGAAGCCCTCTTGGCCTT 1972
Qy 1411 TGTTCACAAATTGATCTCTGTATGCCCTTGGCTTTTCCAGATGCTGCGCCTTC 1470
Db 1973 GGTTCACAACTCTCTCTGCTGCTGACAGTGGCTATTCCAACTATCATCATCTTC 2032
Qy 1471 AAGATGGCCGCGCTGCTCCAGCGGCTACAGACTGCTGGGTCCTGCTACAGGGCCCTAC 1530
Db 2033 AAGTTTCCACCAAGTACATCATATTCTACACGCTGGTCCCAAAACCGCTGCTGGC 2092
Qy 1531 GTCTGTATGAGCTATTCACGAGTACTCAAAATGGTCAATGTGGTAAATGGCATGCTGCA 1590
Db 2093 CTGTTGTGATGATCAGCTCAGCGGCCAGCTGCTATCTGTCTCACTTGGCTGGTGTG 2152
Qy 1591 CGGCTCAGTCCACCCCTGACTGACCCCGATGACCCCAAGATCACAATGTCTCTCTGT 1650
Db 2153 TGGACCCCACTGCTGCTAGGGAATACAGCGCTCCCAATGCTGTATGATGCTGATGC 2212
Qy 1651 AACCCCACTACCGCAACAGCGCTGCTTCAACACAGCGCTGAGCCTGCTGCTCACTG 1710
Db 2213 ACAGAGCAAACTCCCTGCGCTTCTACTGCTTCTCTTCAATGCGCTCTCTCAATC 2272
Qy 1711 GTGGTTTCACTTCCCTACATGAGCAAGAGCTGCCACCACTACACAGAGGCCAAG 1770
Db 2273 AGTGGCTTGGCTGAGCTACTGAGTAAGACTTGGCCAGAGAACTACACAGAGGCCA 2332
Qy 1771 TTCACTACCTCAGATACCTTCTATTTCACTCATCTGCTGCTGCTGCACTTCAATG 1830
Db 2333 TGTGCTACCTTCACTGCTCTTCAACTGCTGCTGATGCTGCTTCAACCAAGGCC 2392
Qy 1831 TCTGCTACAGCGGGGCTGCTGTCACATGCTGACCTGCTGCTGCTCAACCTC 1890
Db 2393 AGCGCTACAGAGGAGTACTGCTGCTGCGCAACATGATGCTGCTGCTGCTGCTG 2452
Qy 1891 CTGGCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1950
Db 2453 AGCAGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2512
Qy 1951 CGCAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2008
Db 2513 CTCAACAGCAGAGCACTTCCAGGCTTCACTTCAAGGACTACAGAGCGCTGCGCT 2570

RESULT 9
AAZ50745
ID AAZ50745 standard; cDNA: 2333 BP.
XX
AC AAZ50745;
XX
DT 31-MAY-2000 (first entry)
XX
DE Human sensory transduction G-protein coupled receptor-B3 cDNA.
XX
KW Human; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
KW sensory cell; taste receptor cell; screen; taste modulator;
KW pharmaceutical; food industry; taste topographic map; tongue; ss.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
FH 1..2333
FT CDS
FT /tag= a
FT /product= "GPCR-B3"
FT /transl_except= (pos:865..866, aa:His)
XX
PN MO200006592-A1.
XX
PD 10-FEB-2000.
XX
PE 27-JUL-1999; 99MO-US17099.
XX
PR 28-JUL-1998; 98US-0094465.
XX

```

```

PA (REGC ) UNIV CALIFORNIA.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Zuker CS, Adler JE, Lindemeier J, Ryba N, Hoon M;
XX WPI: 2000-205451/18.
XX P-PSDB: AA45023.
XX
XX New isolated sensory transduction G-protein coupled receptor, useful
XX for developing products for use in studying and modulating the taste
XX transduction pathway -
XX
XX Claim 5; Pages 79-80; 83pp; English.
XX
XX The present sequence is a cDNA encoding a taste cell specific G-protein
XX coupled receptor, GPCR-B3 which is involved in sensory transduction.
XX This sequence was isolated from the human testis library.
XX GPCR-B3 is specifically expressed in foliate and fungiform cells, with
XX lower expression in circumvallate taste receptor cells of the tongue.
XX The present sequence is used to screen compounds that modulate sensory
XX signalling in taste cells, especially taste modulators useful in
XX pharmaceutical and food industries to customise taste. The sequence
XX can also be used as probe for identifying taste cells and
XX subsets of taste receptor cells such as foliate, fungiform and
XX circumvallate. Such probes are also useful to generate taste
XX topographic maps that elucidate the relationship between the taste
XX cells of the tongue and sensory neurons leading to taste centres
XX in the brain.
XX
XX Sequence 2333 BP; 445 A; 705 C; 646 G; 537 T; 0 other;
XX
XX Query Match 18.7%; Score 376.6; DB 21; Length 2333;
XX Best Local Similarity 50.9%; Pred. No. 1.7e-69;
XX Matches 1036; Conservative 0; Mismatches 954; Indels 47; Gaps 4;
XX
Qy 1 ATCACTTACAGGCGCATGAGTGGTGGAGCAAGTGGCTTGGCTGCTG 60
Db 307 ATTACCTATGGGCGCCAGCAGCAGCAGCTCAGCTGAAAGCGCAATCCCTTCTG 366
Qy 61 CGTACACAGCCAGCGCCAGCAGCAGCAGTGGAGGCTGCTGATGCTGCACTTC 120
Db 367 CGCACCATCCCAATGACAAAGTACAGGTGAGACCTGATGCTGCTGCAAGATTC 426
Qy 121 CGCTGAGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 427 GGGTGGACCTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486
Qy 181 CTGCTTGGCGAGCGGCTGCGCGGCGACATCTGATGCGCTTCCAGAGAGCGTGGCC 240
Db 487 GCACGTGAGAACAGCCCGCTGCTGAGGGGCACTCTGATGCTTCAAGAGCAATGCTCC 546
Qy 241 ACACTGACGCCAACAGACATGACGTGAGAGAGCGCCAGCGCTGCTGCTGCTGCTG 300
Db 547 -----TTCCTGCTCCAGGAGGCGGCTGATGAGAGATGACAGTGCCTCATG 588
Qy 301 GACAACTGACAGCAGACAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 589 CGCACCTGGCCCGAGCGCGGCGGCGACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 648
Qy 361 TACCACTTCTTCAATGAGTGTGCTGCGCAGAACTTCAAGGGCGCGCTGATGCTGCTCC 420
Db 649 AGGTGTTTTCGAGTCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 708
Qy 421 GAGTCTGGGCGCATGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 709 GAAGCTGGGCGCTCTCCAGGCACATCACTGGGCTGCTGCTGCTGCTGCTGCTGCTG 768
Qy 481 TTCTGGGCGCATCAGATCAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 769 GTGCTGGGCGTGGCGCATCAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 828
Qy 541 GGCCACAGCGTGGCGCCAGCCCTCAGAGAGCA-----GCCAGAGTATACCTGCA 595

```


Db 829 TATCCCGGGGAGACAAAGAGGCCCCCTAGGCTTGACAAAGGGCTCTGTCGACAGACA 888
 QY 596 ACCAGAGAGTGGAGAACTGCTGAAAGGCACTTGTCTTCAACACATTCCTCAGGGCTCT 655
 Db 889 ATCAGCTCTGAGAAATGCCAGCTTTCATGCGACACACATGCGCCAGCTCAAGCTT 948
 QY 656 CTGGGAGCGTGTCTACAGCGCTTACCTGCGGTCTATGTCGTGGCCATGCGCTGC 715
 Db 949 TCTCCATGATGTCGCTACACACATACCGATACCGGCTGTATGCGGTGGCCATGCGCTGC 1008
 QY 716 ACAGCGCTCTGCGGTGACAAAAGACACCTGACCAAGAGGGTGTCTACCCCTGGGACG 775
 Db 1009 ACAGCTCTGCGGTGCTGCTGCTGAGCTCTGTTCAGGGGCGGAGCTCACTGCGGACG 1068
 QY 776 TGTCTGAGAGATCTGAAAGGTCAACTTCACTCTCTGAGACACCAATCTCTTCGAC 835
 Db 1069 TTTTGGACAGATCCACAAAGTGTGATTCCTTACACAAAGACACCTGTGGCGTTTATG 1128
 QY 836 CGCAAGGGAGCGTGTCTGCACTTGGAGATTTCCAGTGGCAATGGGACCGGAGCCAGA 895
 Db 1129 ACAACAGAGATCCCTCAGTAGTATTAATATGCTGTGGAGCTGGAATGAGCCAAAT 1188
 QY 896 ATCCCTTCAGAGGCTGCTGCTCTAC-----TACCCCTGAGGAGACGCTGAAGAACA 949
 Db 1189 GGACCTTCAGGCTCTGCTGCTCTCTCCACATGCTCTCACTGCTGAATCAATTAATGAGA 1248
 QY 950 TCAAGACATCTCTGACACACCTCAACACACAGATCCATATGCTCAATGTTTCAAGAGT 1009
 Db 1249 CCAAAATCCAGTGGCAGGAAAGAAACACAGGTGCGTAACTGTGTTCAGCGACT 1308
 QY 1010 GCCAGTCAGGGCAAAAGAAAGCGCTGTGGGATCCACGCTGCTGCTTCAGTGCATCG 1069
 Db 1309 GTCCTGAAGGGCAGCAGAGAGTGTATGCGGTTTCCATCACTGCTGCTTGAAGTGTGTC 1368
 QY 1070 ACTGCTTCCCGGACCTTCTCTCAACACACTGAA-----TGCCGGA 1111
 Db 1369 CTTGTGGGCTGAGACCTTCTCTCAACAGAGCGAGCTTACAGATGCCAGCTTGTGAA 1428
 QY 1112 ATAAAGAGTGTCTTACCAAGATGAGACCTCTGCTTCAAGCGGAGCTGTCTCTG 1171
 Db 1429 CAGAAAGTGGGACCTGAGGAGGAGCCAGACCTGCTTCCGCGACCTGTGTGTTTGG 1488
 QY 1172 AATGGCATGAGGACCCACCATGCTGTGCGCTGTGCGCCCTGGGCTTCTCAGCA 1231
 Db 1489 CTTGCGGTGAGACACCTTGTGGTGTGCGGAGCTTACACACCTGCTGTGCTGTC 1548
 QY 1232 CCTGGCCATCTGTGATTTCTGAGGACCTTCCAGACACCATAGTTCGTGCGCTG 1291
 Db 1549 TGTCTGGGAGCTGCGCTGTTTGGCTGTGACCTAGACACCTGTGTGAGTCAAGCAG 1608
 QY 1292 GGGGCCCCATGCTGCTGATGTGACACTGTGCTGTGGGCAATATGTTGTCCCGG 1351
 Db 1609 GGGGCCCCGTGTCTTCTTATGCTGTGGGCTCTTGGCAGAGGTAATGGAGCCCTATG 1668
 QY 1352 TGTACGTGGGGCCCCCAAGGTCTCCACTGCTGCTGCGCCGAGCCCTTTCCTCT 1411
 Db 1669 GCTTCTTGGGGAACCAAGAGGCTGTGCTGTGCTACGCGAGCCCTTTCGCTTGTG 1728
 QY 1412 GCTTCACAATTTGATGCTGTAATGCGCGCTGCTTCTTCCAGATGCTGCGCGCTCA 1471
 Db 1729 GTTTCACATTTCTCTGCTGCTGAGAGTTCCTCATTTCACTAATCAATCAATCTTCA 1788
 QY 1472 AGATGGCCAGCGCTCCACAGCGCTACAGACTGTGCTGCGCTGCTGAGGCGCTTACG 1531
 Db 1789 AGTTTCCACAAAGTATCTTCAACAGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1848
 QY 1532 TCTCTATGAGCTTATACAGGTACTCAAAATGTCAATTTGTGTAATTTGCAATGCTG 1591
 Db 1849 TGTGTGATGATCAGCTACAGCGCGCCAGCTGTATCTGTCAACTGTGCTGCTGCTG 1908
 QY 1592 GGGCTGAGTCCACCCCGCTGCTGATGCGCCATGACCCCAAGATACCAATTTGTCTCTGTA 1651
 Db 1909 GGAACCCAGCTGCTGTAGGGAATACAGCGCTTCCCGCATCTGTGTGATGCTGAGTGA 1968

QY 1652 ACCCAACTACCCGACAGCGCTGCTTCAACACACCGCTGAGACCTGCTGCTCAGTGG 1711
 Db 1969 CAGAGACCAATCCCTGGCTTCACTGAGCTTCTCTCAATAGGCTCTCTCATCA 2028
 QY 1712 TGGGTTTACGCTTGCCTACATGAGGCAAGAGCTGCCACCACTACAGAGGCCAAT 1771
 Db 2029 GTGCTTGTGCTGACAGTACCTGTGTAAGGACTTGCAGAGATACAGAGGCCAAAT 2088
 QY 1772 TCATCACCCTGAGTACCTTATTTACCTATGCTGCTGCTGCTGCTGCTGCTGCT 1831
 Db 2089 GTGTACCTTACCTGCTGCTTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2148
 QY 1832 CTGCTACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1891
 Db 2149 GCGTCTACAGAGGCAAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2208
 QY 1892 TGGCCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1951
 Db 2209 GCAGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2268
 QY 1952 GCACAGCGCGGCTTCAACAGCATGATCCAGGCTACACCATGAGAGGAGCT 2008
 Db 2269 TCAACAGCAGACACTTCCAGGCTTCCATTCAGGACTACACGAGCGCTGCGCT 2325

RESULT 10
 AD17516
 ID AD17516 standard; cDNA; 2526 BP.
 AC
 AC AD17516;
 DT 10-DEC-2001 (first entry)
 DE Human taste receptor, hT1R1 cDNA coding sequence.
 KW Human; taste-cell-specific G protein-coupled receptor; hT1R1; drug;
 KW genetic modulation; pharmaceutical; taste sensation; food industry;
 KW chemosensory transduction; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..2526
 FT CDS
 FT /tag= a
 FT /product= "Human taste receptor, hT1R1 protein"
 FT /transl_except= (pos:820..823, aa:Phe)
 FT /note= "this codon has an apparent 1 nucleotide insertion
 which alters the reading frame"

WO200166563-A2.
 13-SEP-2001.
 PF 07-MAR-2001; 2001WO-US07265.
 XX
 PF 07-MAR-2000; 2000US-0187546.
 PR 07-APR-2000; 2000US-0195536.
 PR 06-JUN-2000; 2000US-0209840.
 PR 23-JUN-2000; 2000US-0214213.
 PR 17-AUG-2000; 2000US-0226448.
 PR 03-JAN-2001; 2001US-0259227.
 PA (SENO-) SENOMYX INC.
 XX
 PI Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
 XX WPI; 2001-582267/65.
 DR P-PSDB; AAE10372.
 XX
 PT New mammalian taste-cell-specific G protein-coupled receptor
 PT polypeptides for identifying compounds that modulate taste signaling
 PT are useful in food, to modulate the sweet taste of foods or drugs

xx Claim 1; Page 82-83; 119pp; English.

CC The invention relates to mammalian taste-cell-specific G protein-coupled
CC receptors, T1R and their corresponding cDNA molecules. Taste receptors,
CC T1R are useful for screening compounds which are used to activate or
CC modulate chemosensory transduction, such as taste sensation. The
CC identification and isolation of novel taste receptors and taste
CC signalling molecules allow for new methods of chemical and genetic
CC modulation of taste transduction pathways. The taste modulating
CC compounds are useful in pharmaceuticals and food industries to improve
CC the taste of a variety of consumer products, or to block undesirable
CC tastes, e.g., in certain pharmaceuticals. T1R's are also useful in
CC biochemical assay for identifying tastant (T1R) ligands having binding
CC specificity for T1R involved in taste signalling. The present cDNA
CC sequence is human taste-cell-specific G protein-coupled receptor, hT1R1
CC coding sequence.

xx Sequence 2526 BP; 470 A; 770 C; 697 G; 589 T; 0 other;

Query Match 18.3%; Score 368.8; DB 22; Length 2526;
Best Local Similarity 50.9%; Pred. No. 7.7e-68;
Matches 1038; Conservative 0; Mismatches 952; Indels 48; Gaps 5;

Qy 1 ATCACTACAGCCGATCATGAGTGCAGACAAAGTCCGCTTCCGGCTTGTGCG 60
Db 499 ATTAGCTATGCGGCGAGGAGAGAGAGTCAAGTCAAGGAGATGCTTCTTCTG 558
Qy 61 CGTACCAACCCAGGCGGAGCAGACAGTGAAGCCAGTGTGAGTGTGACATTG 120
Db 559 CCGACCATCCCAATGACAAATGACAGGAGGAGACCATGCTGCTGTCGAGAAATTG 618
Qy 121 CGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
Db 619 GGGTGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 678
Qy 181 CTGCTTGGGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
Db 679 GCACTGAGAAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 738
Qy 241 ACATGCAAGCCCAACCAACATGACATGACATGACATGACATGACATGACAT 300
Db 739 -----TTCTCTGCCCCAGGTGGGCGATGATGATGATGATGATGATGAT 780
Qy 301 GACAAGCTGACAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAG 359
Db 781 GCGCAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
Qy 360 GTACCACTTCTCAATGAGAGTGTGCGGCGGCGGCGGCGGCGGCGGCGG 419
Db 841 CAGGGTGTGTTTGAAGTCCGTGTGACCAACCTGACTGAGTGGAGTGGGCG 900
Qy 420 CGAGTCCGAGGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAG 479
Db 901 ACAAGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
Qy 480 CTTCCTGGGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAG 539
Db 961 GGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
Qy 540 GGGCCCAAGAGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 594
Db 1021 TATGCGCGGCGGAGACAAAGAGGCGGCGGCGGCGGCGGCGGCGGCG 1080
Qy 595 AACCAAGAGTGGAGACAGTGGAGAGTGGAGAGTGGAGAGTGGAGAGTGG 654
Db 1081 AATCAAGCTGTGAGAGAAATGCGAAGCTTTCATGAGCAGACAGATGCC 1140
Qy 655 TCTGGGAGAGCTGTGTCTACAGCTGTGTCTGTGTGTGTGTGTGTGTGTGT 714
Db 1141 TTCTCATAGATTTGCGCTACAGAGCATACCGGCGTGTGTGTGTGTGTGT 1200
Qy 715 CACAGGCTCTGCGCTGTGACAAAAGACCTGACCAAGAGGAGTGTGTGTGT 774

Db 1201 CACAGGCTCTGCGCTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Qy 775 CTGCTTGGAGAGATGTGAAAGTCAACTTCACTCTCTCTGAGACCAACCAATCTTGTGAC 834
Db 1261 CTTTGGAGAGATGTGAAAGTCAACTTCACTCTCTCTGAGACCAACCAATCTTGTGAC 1320
Qy 835 CCGCAAGGAGAGT 894
Db 1321 GACAAAGAGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1380
Qy 895 AATCCCTTCAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 948
Db 1381 TGGACCTTACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440
Qy 949 ATCAAGACATCTGTGACACCGTCAACACAGACAGTCTATGTGTGTGTGTGT 1008
Db 1441 ACCAAATTCAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
Qy 1009 TGGCAGTCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1068
Db 1501 TGTCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
Qy 1069 GACTGCTTCCCGGAGCTTCTCTCAACCACTGAA-----TGCCCG 1110
Db 1561 CCGTGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
Qy 1111 AATTAAGAGT 1170
Db 1621 AATTAAGAGT 1680
Qy 1171 GANTGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1230
Db 1681 GCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
Qy 1231 ACCCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1290
Db 1741 CTGCTTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
Qy 1291 GGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1350
Db 1801 GGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
Qy 1351 GTGTAGTGGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1410
Db 1861 GCGTCTTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
Qy 1411 TGCCTTCAATTTGATCTCTGTATGCGGCGGCGGCGGCGGCGGCGGCGGCG 1470
Db 1921 GGTTCACCATCT 1980
Qy 1471 AAGATGGCGAGGCGGCTTCCAGCGGCGGCTTACAGTACTGGGTCCGCTAC 1530
Db 1981 AAGTTCCTCAACCAAGTACTCATATTCACAGCGGCGGCGGCGGCGGCG 2040
Qy 1531 GTCTATGAGCAATTTATCAGGTAATCAAAATGATGATGTGTATTTGGATG 1590
Db 2041 CTGTTTGTATATATCAGTCTACAGCGGCGGCGGCGGCGGCGGCGGCGG 2100
Qy 1591 CGGCTCACTCCACCCCGTACTGACCCGATGACCCGATGACCAATTTCTCTGT 1650
Db 2101 TGGACCCCACT 2160
Qy 1651 AACCCCACTTACCGCAAGCGCTGTCTGTCTCAACACAGCTGTGTGTGTGT 1710
Db 2161 ACAGAGACCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2220
Qy 1711 GTGGGTTTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1770
Db 2221 AATGCTTTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
Qy 1771 TTATCATCCCTGAGATGACCTTCTATTTTCACTCTCTCTCTCTCTCTCT 1830


```

QY 1010 GCCAGTCAGGGCAAAAGAGAGCCTGTGGCATTCCACGTCCTGCTGCTGAGTGCATCG 1069
DB 1515 GTTCGAGAGGGACACACAGGTTGTATGGTTCCACACACTGCTGCTGAGTGCATGCG 1574
QY 1070 ACTGCTTCCCGGACCTTCTCACCACACTGAA-----TGCCCGA 1111
DB 1575 CCGTGAAGCTGGGACACTTCTCAGACGAGTGAAGCTTCACACTGCGACCCCTTGTGAA 1634
QY 1112 ATACAGAGTGTCTTACACAGATGAGACCTCTGCTTCAAGCGGACGACGCTCTCTCG 1171
DB 1635 CACAAGATGAGGCGCCCTGAGAGGAGCTCAGCTCTCTCAGCAGCCAGCTGAGATCTTGG 1694
QY 1172 AATGGCATGAGGACCAACCATGCTGTGGCCCTGCTGCGGCGCCCTGAGCTTCTCGCA 1231
DB 1695 GGTGGCATGAGAACCATCTCTTGTGCTATTGACAGCTAACACGCTATTCCTGCTGTC 1754
QY 1232 CCCCTGGCCATCTGTGTATATTCGTGAGGACATTCACAGACCCATAGTTCGCTCGGCTG 1291
DB 1755 TGATTTGGAGTGTGGGCTGTTGGCGTGTTCACAGCCGCTTGTGAGGTGACGCTG 1814
QY 1292 GGGGCCCCATGTCTTCTGATGCTGACACTGCTGCTGTGGCATACATGCTGCTCCGG 1351
DB 1815 GGGGTAGGCTGTCTCTCATGCTGTGGTCTTGGTAGCTGGGAGTTGACGCTCTACCA 1874
QY 1352 TGTACGTGGGGCGCCCAAGGTCCTCACACGTCGCTGCGGCGCCCTCTTCCCTCT 1411
DB 1875 GCTTCTTGGGAAAGCCACAGGTGCGCGGTGCTGCTGCTGACGCCCCCTTCTTCTCTCG 1934
QY 1412 GCTTCAACATTTGATCTCTGATATCCCTGATATCCCTGCTTCTTTCAGATGCTGCGCTTCA 1471
DB 1935 GGTTCGATTTTCTCTCTCTCTGATGATATCCGCTCTTCAACATGCTATCATCTTCA 1994
QY 1472 AGATGGCAGACGCTTCCACAGGCTACAGCTACTGGGTCCGCTACCAAGGGCCCTACG 1531
DB 1995 AGTTTCTTACCAAGTACCCACATTCACACACTTGGGCGCCAAACCAATGCTGCCGAA 2054
QY 1532 TCTCTATGCGATTATACAGGTAAGTCAAAATGTCATTTGATTTGGATGGATGGTCAC 1591
DB 2055 TATTCGATATTGATGAGTCCACAGGTCATTTGTTCTCTCTCTCAGCTGCTTGCATATG 2114
QY 1592 GGCCTGATGCCACCCCGGTAAGTACCCGATGACCCCAAGATCACAATTTGCTCTCTGTA 1651
DB 2115 GGAACCCACAGGCGCCACAGGAGTACAGCGCTTCCCAATCGGTGATTTGAGTGCAT 2174
QY 1652 ACCCCAACTACCGCAGACGCTGCTGTCAACACAGCCTGAGCTGCTGCTCACTGG 1711
DB 2175 CAGAGGTCAACTCTGTGGGCTTCTGTGCTTTCGCACACAACATCTCTCTCCATCA 2234
QY 1712 TGGGTTTACGCTTCGCTTACATGGGCAAGAGCTGCCACACCACTACATACAGAGGCAAT 1771
DB 2235 GCACCTTGTCTGACAGTACTGAGTGAAGTACGCGGAGAACTATACAGAAAGCCAAAT 2294
QY 1772 TCATCAACCTCAGCATGATGACTTCTTATTTACCTCATCCGCTCTGACCTTCATGT 1831
DB 2295 GTGTCAACCTTCAACCTGCTCTCCTCATCTGATTCGTTGATGCTTCTTCCATATGCA 2354
QY 1832 CTGCTTACAGCGGGGTGCTGTGTACCATCTGGAACCTTGTGTGATCTGTCTCAACTCC 1891
DB 2355 GCATTTTACAGGCGGACCTACCTACCCGCGGCTCATGTGCTGGAAGGCTGCGCACTCTGA 2414
QY 1892 TGCCCATTCAGGCTTGGCTTACTTGGGCCCCAAGTCAATGCTATCTTCTTACCCGAGG 1951
DB 2415 GTGGCGGCTTACAGCGGCTATTTCTCTCCCTTAATGCTACGATGATTTCTGCGCTCCAGAAC 2474
QY 1952 GCAACAGCGCCGCTACTTCAACAGCATGATCCAGGGCTTACACATGAGAGGAGAC 2007
DB 2475 TCAACAACAGAACACTTTCAGGCTTCATTCACAGGACTACAGAGGCGCTGGGCGC 2530

```

RESULT 12
 AAS46935
 ID AAS46935 standard; cDNA; 697 BP.

```

XX AC AAS46935;
XX DT 18-DEC-2001 (first entry)
XX DE Human G protein-coupled receptor (GPCR) cDNA #117.
XX XX
XX KW Human; G protein-coupled receptor; GPCR; mental disorder; schizophrenia;
XX KW neurological disorder; metabolic disorder; cancer; rheumatoid arthritis;
XX KW thyroid disorder; neurodegenerative disorder; cardiovascular disorder;
XX KW renal failure; autoimmune disorder; hyperproliferative disorder; HIV; ss;
XX KW human immunodeficiency virus; viral infection; neuroprotective;
XX KW immunostimulant; neuroleptic; nootropic; tranquiliser; antidepressant;
XX KW anorectic; gene therapy.
XX OS Homo sapiens.
XX PN WO200168858-A2.
XX PD 20-SEP-2001.
XX PF 16-MAR-2001; 2001WO-0508456.
XX PR 16-MAR-2000; 2000US-187783P.
XX PR 16-MAR-2000; 2000US-189907P.
XX PR 16-MAR-2000; 2000US-189917P.
XX PR 16-MAR-2000; 2000US-189918P.
XX PR 16-MAR-2000; 2000US-189960P.
XX PR 29-MAR-2000; 2000US-192155P.
XX PR 29-MAR-2000; 2000US-192234P.
XX PR 29-MAR-2000; 2000US-192830P.
XX PR 29-MAR-2000; 2000US-192916P.
XX PR 29-MAR-2000; 2000US-192933P.
XX PR 29-MAR-2000; 2000US-192935P.
XX PR 29-MAR-2000; 2000US-192945P.
XX XX
XX PA (PMAA ) PHARMACIA & UPJOHN CO.
XX PI
XX PI Vogel1 G;
XX DR WPI; 2001-607458/69.
XX DR P-PSDB; AAU29496.
XX XX
XX PT Nucleic acid encoding G protein-coupled receptors, useful for the
XX PT prevention, diagnosis and treatment of mental disorders -
XX PS Claim 4; Page 97; 274pp; English.
XX XX
XX CC Sequences AAS46819-AAS46946 represent cDNA molecules encoding human G
XX CC protein-coupled receptor (GPCR) polypeptides. The protein and DNA
XX CC sequences of the invention can be used to identify compounds which bind
XX CC to GPCR polypeptides and in screening for compounds that modulate GPCR
XX CC activity. By screening a human subject for the presence of mutations in
XX CC GPCR DNA, a GPCR-related disorder or a genetic predisposition can be
XX CC diagnosed. The sequences can also be used for treatment and prevention of
XX CC mental disorders such as schizophrenia, neurological disorders such as
XX CC manic depression, metabolic disorders such as obesity, cancer, rheumatoid
XX CC arthritis, thyroid disorders such as myxoedema, neurodegenerative
XX CC disorders such as Parkinson's disease, cardiovascular disorders such as
XX CC atherosclerosis, renal failure, autoimmune disorders, hyperproliferative
XX CC disorders such as psoriasis and viral infections such as those caused by
XX CC HIV.
XX XX
XX SQ Sequence 697 BP; 130 A; 177 C; 211 G; 179 T; 0 other;
XX XX
QY Query Match 12.8%; Score 258; DB 22; Length 697;
QY Best Local Similarity 100.0%; Pred. No. 9, 6e-45;
QY Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1105 TGCCCGAATTAACGAGTGTCTTACAGAGTGAAGCTCTCTGCTTCAAGCGGACGCTGTC 1164
QY 440 TGCCCGAATTAACGAGTGTCTTACAGAGTGAAGCTCTCTGCTTCAAGCGGACGCTGTC 499

```

QY 1165 TTCTGGAATGGCAGAGGACACCCACCATCGTGTGGCCCTGCTGCGCCCGCTGGGCTTC 1224
 |||||||
 Db 500 TTCCGGAATGGCAGAGGACACCCACCATCGTGTGGCCCGCTGGGCTTC 559
 |||||||
 QY 1225 CTCAGACCCCTGGCATCTCTGTATATTTCTGGAGGACCTTCAGACACCCATAGTTGC 1284
 |||||||
 Db 560 CTCAGACCCCTGGCATCTCTGTATATTTCTGGAGGACCTTCAGACACCCATAGTTGC 619
 |||||||
 QY 1285 TCGGCTGGGGGCCCATGTCCTCTGATGCTGACACGCTGCTGTGGCATACATGCTG 1344
 |||||||
 Db 620 TCGGCTGGGGGCCCATGTCCTCTGATGCTGACACGCTGCTGTGGCATACATGCTG 679
 |||||||
 QY 1345 GTCCCGGTGTACGTGGG 1362
 |||||||
 Db 680 GTCCCGGTGTACGTGGG 697
 |||||||

RESULT 13
 ABR81713
 ID ABR81713 standard; DNA; 697 BP.
 AC ABR81713;
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE cDNA encoding novel G protein coupled receptor (nGPCR-x) #117.
 XX
 KW G protein coupled receptor; nGPCR-x; immune response; thyroid disorder;
 KW mental disorder; thyrotoxicosis; myxoedema; inflammatory condition;
 KW Crohn's disease; cell differentiation; homeostasis; rheumatoid arthritis;
 KW renal failure; autoimmune disorder; movement disorder; CNS disorder;
 KW viral infection; human immunodeficiency virus; HIV; metabolic disorder;
 KW cardiovascular disorder; diabetes; obesity; anorexia; cardiomyopathy;
 KW proliferative disease; cancer; psoriasis; lung cancer; hormonal disorder;
 KW sexual dysfunction; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN US2002058306-A1.
 XX
 PD 16-MAY-2002.
 XX
 PF 16-MAR-2001; 2001US-0811284.
 XX
 PR 16-MAR-2000; 2000US-189783P.
 PR 16-MAR-2000; 2000US-189907P.
 PR 16-MAR-2000; 2000US-189917P.
 PR 16-MAR-2000; 2000US-189918P.
 PR 16-MAR-2000; 2000US-189960P.
 PR 24-MAR-2000; 2000US-192153P.
 PR 27-MAR-2000; 2000US-192234P.
 PR 29-MAR-2000; 2000US-192830P.
 PR 29-MAR-2000; 2000US-192945P.
 PR 29-MAR-2000; 2000US-192915P.
 PR 29-MAR-2000; 2000US-192923P.
 PR 29-MAR-2000; 2000US-192830P.
 PR 29-MAR-2000; 2000US-192945P.
 PR 29-MAR-2000; 2000US-192945P.
 PR 29-MAR-2000; 2000US-192945P.
 PA (VOGE/) VOGELI G.
 XX
 PI Vogel G;
 XX
 DR WPI; 2002-434856/46.
 DR P-PSDB; ABR60784.
 XX
 XX
 PT New isolated nucleic acid encoding a G protein coupled receptor for
 PT producing the receptor which can induce an immune response in a mammal
 PT
 XX

PS Claim 4; Page 75-76; 216pp; English.
 XX
 CC The invention describes an isolated nucleic acid (I) comprising a
 CC sequence encoding a portion of a G protein coupled receptor (nGPCR-x).
 CC (I) is used to produce a recombinant nGPCR-x polypeptide. A polypeptide
 CC encoded by (I) is used to induce an immune response in a mammal. nGPCR-x
 CC is used to identify a compound that binds to it and/or modulates its
 CC activity. (I) is used to identify animal homologues of nGPCR-x. (I) can
 CC be used to diagnose a human subject as having a brain or genetic
 CC predisposition disorder, such as a mental disorder. (I) is used to
 CC screen for an nGPCR-x related disorder including thyroid disorders (e.g.
 CC thyrotoxicosis, myxoedema), renal failure, inflammatory conditions (e.g.
 CC Crohn's disease), diseases related to cell differentiation and
 CC homeostasis, rheumatoid arthritis, autoimmune disorders, movement
 CC disorders, CNS disorders, viral infections (e.g. Human immunodeficiency
 CC virus), metabolic and cardiovascular disorders (e.g. diabetes, obesity,
 CC anorexia, cardiomyopathies), proliferative diseases and cancers (e.g.
 CC psoriasis, lung cancer), hormonal disorders, sexual dysfunction and
 CC hereditary mental disorders in a human patient. A host cell comprising
 CC (I) is used to screen for a modulator of nGPCR-x activity. nGPCR-x is
 CC used to identify compounds that can treat mental disorders. The
 CC polypeptide encoded by (I) is used to purify a G protein from a sample.
 CC This sequence encodes a novel G protein coupled receptor (nGPCR-x)
 CC protein described in the invention.
 CC
 CC
 SQ Sequence 697 BP; 130 A; 177 C; 211 G; 179 T; 0 other;

Query Match 12.88; Score 258; DB 24; Length 697;
 Best Local Similarity 100.0%; Pred. No. 9; 6e-45;
 Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1105 TGCCGGAATAGAGAGTGTCTTACAGAGTGAACCTCTCTCAAGCGGACGTGGTC 1164
 |||||||
 Db 440 TGCCGGAATAGAGAGTGTCTTACAGAGTGAACCTCTCTCAAGCGGACGTGGTC 499
 |||||||
 QY 1165 TTCTGGAATGGCAGAGGACACCCACCATCGTGTGGCCCTGCTGCGCCCGCTGGGCTTC 1224
 |||||||
 Db 500 TTCCGGAATGGCAGAGGACACCCACCATCGTGTGGCCCGCTGGGCTTC 559
 |||||||
 QY 1225 CTCAGACCCCTGGCATCTCTGTATATTTCTGGAGGACCTTCAGACGCGCATAGTTGC 1284
 |||||||
 Db 560 CTCAGACCCCTGGCATCTCTGTATATTTCTGGAGGACCTTCAGACGCGCATAGTTGC 619
 |||||||
 QY 1285 TCGGCTGGGGGCCCATGTCCTCTGATGCTGACACGCTGCTGTGGCATACATGCTG 1344
 |||||||
 Db 620 TCGGCTGGGGGCCCATGTCCTCTGATGCTGACACGCTGCTGTGGCATACATGCTG 679
 |||||||
 QY 1345 GTCCCGGTGTACGTGGG 1362
 |||||||
 Db 680 GTCCCGGTGTACGTGGG 697
 |||||||

RESULT 14
 ABR47345
 ID ABR47345 standard; CDNA; 2559 BP.
 XX
 AC ABR47345;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE cDNA sequence encoding human AXOR79.
 XX
 KW AXOR79; human; seven transmembrane; G-protein coupled; receptor; ss;
 KW gene; bacterial; viral; fungal; infection; protozoal; manic depression;
 KW human immunodeficiency virus; HIV; cancer; diabetes; anorexia; bulimia;
 KW Parkinson's disease; heart failure; asthma; allergy; osteoporosis;
 KW neurological disorder; schizophrenia; Huntington's disease.
 XX
 OS Homo sapiens.
 XX
 FH key Location/Qualifiers
 FT CDS 1..2559
 FT /*tag= a

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OK nucleic - nucleic search, using sw model

Run on: May 23, 2003, 12:30:29 ; Search time 2803.62 Seconds
(without alignments)
11611.043 Million cell updates/sec

Title: US-09-927-315-12

Perfect score: 2010

Sequence: 1 atcacctacagcgcacatcag.....acacatcagagagagactag 2010

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estbun:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	293.4	14.6	525	17	AZ483105	AZ483105 IM0308M03
2	260.2	12.9	530	17	AZ342774	AZ342774 IM0076F07
3	154.2	7.7	1014	17	CNS033XY	AI226735 Tetradon
4	114	5.7	979	17	CNS030LO	AI223281 Tetradon
5	97.8	4.9	973	17	CNS02BUD	AI190335 Tetradon
6	96.4	4.8	1038	17	CNS0506R	AI315180 Tetradon

C 7	95.6	4.8	562	9	AI742401	AI742401 wq40e02.x
C 8	92.8	4.6	938	13	BI462667	BI462667 603202107
C 9	90.2	4.5	965	17	CNS03DIP	AI236834 Tetradon
C 10	82.8	4.1	298	9	AI562167	AI562167 vw73d10.x
C 11	82.2	4.1	666	10	BB618551	BB618551 BB618551
C 12	80.8	4.0	816	17	AZ535744	AZ535744 ENTC025TR
C 13	80.6	4.0	496	9	AA853967	AA853967 aj51e10.s
C 14	76.6	3.8	890	17	BH146886	BH146886 ENTPK48TF
C 15	76.4	3.8	577	17	AZ343139	AZ343139 IM0076F07
C 16	76.4	3.8	877	17	AZ531291	AZ531291 ENTBQ34TR
C 17	76.4	3.8	912	17	AZ551092	AZ551092 ENTFJ22TF
C 18	76.2	3.7	908	17	AZ548467	AZ548467 ENTEK30TR
C 19	74.6	3.7	1041	17	CNS012X1	AI174862 Tetradon
C 20	73.6	3.7	525	9	AI390993	AI390993 mb98d01.y
C 21	73.6	3.7	891	17	AZ683582	AZ683582 ENTRK47TR
C 22	73.2	3.6	829	17	CNS04565	AI274982 Tetradon
C 23	73	3.6	880	17	AZ529191	AI274981 ENTBV68TR
C 24	72.6	3.6	906	17	BH153606	BH153606 ENTSB3TF
C 25	71.6	3.6	577	17	AZ744141	AZ744141 RPCI-24-1
C 26	70.6	3.5	843	17	AZ551618	AZ551618 ENTDV54TR
C 27	69.4	3.5	421	9	AI415100	AI415100 mb98d01.x
C 28	69	3.4	931	17	BH160272	BH160272 ENTOV49TR
C 29	68.8	3.4	976	17	BH149983	BH149983 ENTOD93TF
C 30	68.2	3.4	890	17	AZ530768	AZ530768 ENTBH54TF
C 31	68	3.4	750	14	BQ178590	BQ178590 UI-M-EVO-
C 32	67.4	3.4	626	17	AZ994369	AZ994369 2M02791L8
C 33	67.2	3.3	849	17	AZ546009	AZ546009 ENTFW53TF
C 34	67	3.3	900	17	AZ549880	AZ549880 ENTD94TF
C 35	66	3.3	717	17	AQ095343	AQ095343 RPCI-23-3
C 36	65.2	3.2	1078	17	CNS054KQ	AI320867 Tetradon
C 37	64.4	3.2	888	17	AZ528430	AZ528430 ENTC024TR
C 38	64.2	3.2	445	13	BM167469	BM167469 EST569992
C 39	63.8	3.2	742	10	AV674403	AV674403 AV674403
C 40	63.4	3.2	439	10	AA496772	AA496772 ESTPM077
C 41	63.4	3.2	905	17	AZ530256	AZ530256 ENTEV38TR
C 42	63.2	3.1	1023	17	CNS05805	AI326174 Tetradon
C 43	63	3.1	598	13	BM170666	BM170666 EST573189
C 44	63	3.1	605	13	BM163520	BM163520 EST566043
C 45	63	3.1	645	13	BM165350	BM165350 EST567873

ALIGNMENTS

RESULT 1
LOCUS AZ483105 525 bp DNA linear GSS 05-OCT-2000
DEFINITION IM0308M03R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0308M03 R, DNA sequence.
ACCESSION AZ483105
VERSION AZ483105.1 GI:10646785
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclauognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 525)
AUTHORS Dunn,D., Moyagi,A., Barber,M., Beecorn,T., Duval,B., Hamli,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
'M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausen,A.
and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL COMMENT
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

Plate: 0308 row: M column: 03
 Seq primer: CACACAGGAAACAGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 525.
 Location/Qualifiers

FEATURES

source

1.525
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUCG1M0308M03"
 /clone_1lb="Mouse 10kb plasmid UUCG1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
 ORIGIN

68 a 165 c 151 g 141 t

Query Match 14.6% Score 293.4; DB 17; Length 525;
 Best Local Similarity 77.8%; Pred. No. 3.8e-51;
 Matches 354; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 1105 TGCCGATTAACGAGTGGTCTTACACAGTGGACCTCTGCTTCAAGCGGACAGTGGTC 1164
 Db 70 TGCCGGGTTCCATGTGTGTTTACAGAACATCGTGTCTTCAAGCGGCGCTGGCC 129
 QY 1165 TTCTGGAATGGCATGAGGACCCACCATCGCTGTGGCCGCGCTGGCTTC 1224
 Db 130 TTCTGGAATGGCATGAGGACCCACCATCGCTGTGGCCGCGCTGGCTTC 189
 QY 1225 CTCGACACCCGTCGCTGGTGAATTCGAGGACCTTCAGACCCCATAGTGGC 1284
 Db 190 ATCAGTACGCTGGCATTTCTGCTTCTTGGAGACATTCACAGCCCATGTGGC 249
 QY 1285 TCGGCTGGGGCCCATGCTGCTTCTGACATGCTGCTGCTGCTGCTGCTG 1344
 Db 250 TCGGCTGGGGCCCATGCTGCTTCTGACATGCTGCTGCTGCTGCTGCTG 309
 QY 1345 GTCCGGTGTACGTGGGGCCGCAAGGTCTCACTGCTTCCGCGGCGCTCTT 1404
 Db 310 GTCCGGTGTACGTGGGGCCGCAAGGTCTCTGCTGCTGCTGCTGCTGCT 369
 QY 1405 CCCCTGCTTCAATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1464
 Db 370 ACCGTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 429
 QY 1465 GCGTTTCAAGTGGCCGCTTCCACAGGCGCTACAGCTAGTGGTCCGCTACAGGG 1524
 Db 430 GTCTTCAAGTGGCCGCTTCCACAGGCGCTACAGCTAGTGGTCCGCTACAGGG 489
 QY 1525 CCTACGCTCTATGCAATTAACAGGTAATA 1559
 Db 490 CCTACGCTCTATGCAATTAACAGGTAATA 524

RESULT 2

AZ342774 530 bp DNA linear GSS 29-SEP-2000
 LOCUS 1M0076F07F Mouse 10kb plasmid UUCG1M library Mus musculus genomic
 DEFINITION clone UUCG1M0076F07 F, DNA sequence.

ACCESSION AZ342774
 VERSION AZ342774.1 GI:10420346
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia, Sciurognathi: Muridae: Murinae; Mus.
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingy,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SIC, UT 84112, USA
 Tel.: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu

Insert length: 10000 Std Error: 0.00
 Plate: 0076 row: F column: 07
 Seq primer: CATTGTAAACGACGCGCCACT
 Class: plasmid ends
 High quality sequence stop: 530.
 Location/Qualifiers

FEATURES

source

1.530
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUCG1M0076F07"
 /clone_1lb="Mouse 10kb plasmid UUCG1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (g114732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 113 a 175 c 136 g 105 t 1 others

ORIGIN

Query Match 12.9% Score 260.2; DB 17; Length 530;
 Best Local Similarity 74.2%; Pred. No. 3.4e-44;
 Matches 328; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 357 CTTGTACACTTCTTCAATGAGTGTGCGCCAGACCTTCAAGCGGCGCTGTGATGCG 416
 Db 1 CTTGTACACTTCTTCAATGAGTGTGCGCGTGAATTCACAGCGCTTGTGATGCG 60
 QY 417 CTCGAGTCTGGGCGCATCGACCCGCTCTGACACACCTCAAGAGCTGGCCACTGGG 476
 Db 1 CTCGAGTCTGGGCGCATCGACCCGCTCTGACACACCTCAAGAGCTGGCCACTGGG 476

Db 61 CTGTAGTCTGTGGGCGATGAGACCTGTCTACACAACTCAAGAGTGGCCACAGGG 120
 QY 477 CACCTTCTGGGCAATACCAATCCAGAGCGTCCCATCCGGGCTTCACTGAGTCCGGA 536
 Db 121 CACTTCTCTGGGCGGCGACATCCAGAGGCTGTCATCCCTGGCTTCAAGCCAGTCCGAGT 180
 QY 537 GTGGGGCCCAAGGCTGGGCGGCGCCACCCCTCAGACGAGCCAGCACTTACTCTGAA 596
 Db 181 GCGCCACACAAAGCCAGATATCCCATGCTTACGACAGACCGCTGAGAGCTACTCTGAA 240
 QY 597 CCAGAGTGGCAACACTCTGTAAGCGCCACTTGTCTTCAACACACTTCTCAGGCTTC 656
 Db 241 CCAGAGTCTGACCGCTCATGATGACATCAGCGAGTCTTAAACAGCTTCTCATGCTTTC 300
 QY 657 TGGGAGGCTGTCTCTACAGGCTGTACTGTGCGGTCTATGCTGTGGCCCATGCCCTGCA 716
 Db 301 GGGGAGGCTGTGTCTCTACAGTGTACTGCGCGCTCTACGCGGAGCCACACCTCTCA 360
 QY 717 CAGGCTCTGCGGTGTGCAAAAGCACTGACACAGAGGCTGTCTACCCCTGGCAGCT 776
 Db 361 CAGACTCTCTCAGCTGCAACAGATCCGCTGCAACCAAGCAAAATCTATTCATGGCAGGT 420
 QY 777 GCTTGAAGAGATCTGAAGGTC 798
 Db 421 GAGGGCCCTAANCTGCGAGGGAC 442

RESULT 3
 CNS033XY/C
 LOCUS
 DEFINITION
 209N08 of library G from Tetraodon nigroviridis, genomic survey
 sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 AL226735.1 GI:7885667
 GSS: genome survey sequence.
 Tetraodon nigroviridis.
 Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Tetraodontidae; Tetraodon.
 1 (bases 1 to 1014)
 Roest-Crolius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Bernot, A., Fitzmes, C., Winkler, P., Brothier, P., Quetier, F.,
 Saurin, W., and Weissenbach, J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence
 Unpublished

JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 2 (bases 1 to 1014)
 Roest-Crolius, H., Jallion, O., Dasilva, C., Fitzmes, C., Fisher, C.,
 Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
 Weissenbach, J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Unpublished
 3 (bases 1 to 1014)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000)
 This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
 SOURCE
 1. 1014
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="209N08"
 /clone_1tb="G"
 /note="Genoscope sequence ID : C0AG209D04LP1-end : T7"
 BASE COUNT
 283 a 209 c 322 g 196 t 4 others

Query Match 7.78; Score 154.2; DB 17; Length 1014;
 Best Local Similarity 50.08; Pred. No. 7e-22;
 Matches 435; Conservative 1; Mismatches 430; Indels 4; Gaps 2;

QY 1105 TGGCGAATTAAGAGTGGTCTTACAGATGAGACCTCTCTTCAACGGCAGCTGTGCT 1164
 Db 875 TGTGAGAGACACAGATGCTTCAACAGCTGGAAGACACTCTCTCTCCACAGGAG 816
 QY 1165 TTCTGGAATGGAGAGGCAACCCACATCGCTGTGGCCCTGCG--TGGCCGCTTGGGCT 1223
 Db 815 TGGGAGCGCTTACAGAGACCCCGGCGGTGTGGATCTCTGGCGGGGCTGCTGCTGT 756
 QY 1224 CCTACAGACCTTGGGCACTCTGTGATATTTTGGAGGACACTTCCAGACACCATAGTTGC 1283
 Db 755 GGGCTTACCCCGACAGATGCTGTCTTGTGGCGCTCACTACACACAGCGGAGTCCG 696
 QY 1284 CTGGGCTGGGGGCCCATGTGCTTCTGATGCTGACATGCTGCTGTGGCATATAGGT 1343
 Db 695 ATCTGCGGGGAGACCATGCTGCTTCTGATCTTAGGCTGCTCAGCTCTGTAGCATAG 636
 QY 1344 GGTCCGGGTGTACGTGGGCGCCGCAAGTCTCCACCTGCTTCCGCGCAGGCTCTT 1403
 Db 635 CGTCTCTTTTACTTTGAGAGACGAGCTGAGAGCTTTGTATCTGAGGTTATGCCCTT 576
 QY 1404 TCCCTCTGTCTTACATTTGATCTCTGTATGCGCGGTCTTCCAGATGCTGTG 1463
 Db 575 TCTTTTATTCATACGCTGTCTGTCTGTGCTTGTGCTGCTTCCGCTCTTTCAGATGCTAT 516
 QY 1464 GGCCTCAAGATGGCGACCGCTTCCGCGGCGCTACAGCTACAGGGGCGCTTCCAGAG 1523
 Db 515 CATTTTAAATATAGCGCGCAAGTTCGCCGAGTCCACAGCTGTGTGAGATGAAATACAGG 456
 QY 1524 GCCCTACGCTCTATGGCATTTATACAGGTACTCAAAATGTCTATTTGTATTTGGCAT 1583
 Db 455 CCAAGGCTGTGTATCTCATGAGAGCTTTGTCTTCCAGCAGTGTGATGTGATCGGCTT 396
 QY 1584 GCTGGCAGGCGCTCAGTCCACCCCGGTACAGAGCCCGATGACCCCAAGATACAAATGT 1643
 Db 395 CTGTGCAACCCACCTGCTTACATGAGCTCTCTCTCTTCCATGAGCAAAATCATCTT 336
 QY 1644 CTCTGTATCCCACTACAGCTGCTGCTGCTTCAACAGCCGCTGAGCTGCTCT 1703
 Db 335 GGGCTGTACGTA--ATCTCAATATGGATGATGAGCGATTTTGTCTTATTTT 279
 QY 1704 CTCACTGTGGTTTACGCTTGTGCTTACATGAGGCAAGAGTGGCCACCAATACAGCA 1763
 Db 278 GTGATCTCTGCTTACTTCTCTCATATGAGGGAAGACCTCCGGAAGAACTACACAGCA 219
 QY 1764 GGGCAAGTTTACACCCCTCAGATGACCTTATTTTCACTGATCCGCTCCCTCGAC 1823
 Db 218 GGGCAAGCCATACCTTCTGCTGCTTCTGCTGATCTCTCAGCTGATCATATTTGCCAC 159
 QY 1824 CTTCATGTCTCTACAGCGGGGTGCTGTGACATCTGTGACCTTGTGTACATGCT 1883
 Db 158 CGGTTATGTCTTACAGCGGGAATATACATCAACCTCAAGCGGCGGTGCTC 99
 QY 1884 CAACCTCTGGCCATCAGCTGGGCTTACTTGGCCCAAGTGTATGATCTCTTCTTA 1943
 Db 98 CAGGCGCTACTGCTTCTTCTGTTGATTTCTCCCAAAATGATATCATCATCTTCA 39
 QY 1944 CCGGAGGCGCAACAGCGCCGCTTACTTCA 1973
 Db 38 GCTTCACAGAACCCAGAAATCTTCA 9

RESULT 4
 CNS030L/C
 LOCUS
 DEFINITION
 184L20 of library G from Tetraodon nigroviridis, genomic survey
 sequence.
 ACCESSION
 AL222381

KEYWORD	AL222381.1	GI:7881200
VERSION	GSS: genome survey sequence.	
SOURCE	Tetraodon nigroviridis.	
ORGANISM	Tetraodon nigroviridis	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.	
TITLE	1 (bases 1 to 979)	
JOURNAL	Roeest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fitzmes,C., Wincker,P., Brothier,P., Queller,F., Saurin,W. and Weissenbach,J.	
REFERENCE	2 (bases 1 to 979)	
AUTHORS	Roeest-Crollius,H., Jallion,O., Dasilva,C., Fitzmes,C., Fisher,C., Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and Weissenbach,J.	
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis	
JOURNAL	Unpublished	
REFERENCE	3 (bases 1 to 979)	
AUTHORS	Genoscope.	
JOURNAL	Direct Submission	
COMMENT	Submitted (12-APR-2000) This sequence is a single read and was generated as part of a large scale clone-and sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon .	
FEATURES	Location/Qualifiers	
source	1..979	
	/organism="Tetraodon nigroviridis"	
	/db_xref="taxon:99883"	
	/clone="184L20"	
	/clone_lib="G"	
	/note="Genoscope sequence ID : C0AG184DF10SP1-end : PUC-Or1"	
BASE COUNT	227 a 230 c 299 g 198 t 25 others	
ORIGIN		
Query Match	5.7%; Score 114; DB 17; Length 979;	
Best Local Similarity	49.4%; Pred. No. 1.8e-13;	
Matches 339; Conservative 12; Mismatches 332; Indels 3; Gaps 3;		
OY	1143 CTGCTTCAAGCGGCGAGCTGGTCTTCTCGTAATGGCATGAGGACCAACCACTGCTGGC	1202
DB	669 CTKTGTGAGAGAGACTCGGCTGCACTGTGATGGACCACTCTATGCTGGCTCTCT	630
OY	1203 CCTGCTGCGCGCCCTGGGCTTCTTCAGCACCCTGGCCATCTCGTATATCTGGAGCA	1262
DB	629 TTTCTTCTGTGCTGTCTCTTCATGACCTCCAGCTCGGCAAGTATCTCTCTCA	570
OY	1263 CTTCAGACACCCATAGTTCGCTCGGCTGGGGGGGCCCATATGCTTCTGATGCTGACCT	1322
DB	569 CATTAACACGCGCTGGGCCAAGTCCGCCGGAGGCGACCTGCCCTTAAAGCTGGCAGC	510
OY	1333 GCTGCTGTGCAATGATGATGTCGCCGCTAGCTGGGGGCGGCCAAGGATCTCCACTG	1382
DB	509 TCTAACACAGCAGCAGTACGAGCTCTMTGTCACCTTTGGCAGCCGCTCCCGCTGGCCAG	450
OY	1383 CCTGCGCGCAAGGCGCTCTTCCCTGCTGCTTCAACAATTCATCTCTG-TAIGCGG	1441
DB	449 CAAGCTCAAGCAGCTCATTCACGTTTCAAGCTTCACTGTGCTTCCGCTCAATCGCTG	390
OY	1442 TGCGTTTTCACAGTCTGTGCGCTTCAAGATGGCAGCGGCTTCCACAGCGCTTACA	1501
DB	389 TGCGTCTGCAACAGCTCATTCATTTTAAATTCGCTCGCGCTGC-CGCGAGCTTACA	331
OY	1502 GCTACTGTGCTCGCTACAGGAGGCGCTGATGCTATGATGATATATCAGGATACGACAA	1561
DB	330 ACARGTGGGCCAAAACCATGGGCGCTCAGGCGCCACATTTTATGCACTGTCCGCAATCC	271

Qy	1562	TGCTATTTGGTAAATTGSCATGCTGGGACGGCCCTCACTCCACACCCCGTACTGACCCCG	1621
Db	270	TGTGGCTCTCTGTCTGGCCGTAAGCCGTC - GGGCCCCCCAGCCCTCCCAARRATCTTATTT	212
Qy	1622	ATGACCCCAAGATCAAAATTGTCTTCCTGTAAACCCCACTACGCAACAGACCTGCTTTCA	1681
Db	211	TCMAACCAAAAGCATCGMGTCTGTAAGMCCAGAACACCTCTCGCCGCGTTCTCTGTGG	152
Qy	1682	ACACGAGCCTGACCTGCTGCTCTCAAGTGGGGTTTCAAGTCTGCCCTCAACATGGGCAAG	1741
Db	151	AGCTCTGTAAAGCTCACTGTCTGTGAAGCGCGTTTGCTTGTTCATGCAACATGGGCAAR	92
Qy	1742	AGCTGCCCAACCAATCAACAAGAGGCCAAGTTTCATCAACCTCAGCATGACTTCTTATTTCA	1801
Db	91	ACCTCTCGGCAACATCAACAAGGAGGCAAGTGGGTCAACCTCAGCTGATGGAGCAATRA	32
Qy	1802	CCTCATCGCTCTCCCTCTGCACCTTC	1827
Db	31	TCTCTCGATTBAGCTTCTTTCACCGTTC	6

RESULT 5	
CNS02BUU	
LOCUS	973 bp DNA linear GSS 12-MAY-2000
DEFINITION	Tetradon nigroviridis genome survey sequence pUC-ori end of clone 254011 of library G from Tetradon nigroviridis, genomic survey sequence.
ACCESSION	AL190335
VERSION	AL190335.1 GI:7828439
KEYWORDS	GSS; genome survey sequence.
SOURCE	Tetradon nigroviridis.
ORGANISM	Tetradon nigroviridis
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
AUTHORS	1 (bases 1 to 973) Roest-Crollius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fitzames,C., Wincker,P., Brottier,P., Queller,F., Saurin,W. and Weissenbach,J.
TITLE	Human gene number estimate provided by genome wide analysis using Tetradon nigroviridis DNA sequence
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 973) Roest-Crollius,H., Jalllon,O., Dasilva,C., Fitzames,C., Fisher,C., Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and Weissenbach,J.
AUTHORS	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetradon nigroviridis
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 973) Genoscope.
AUTHORS	Direct Submission
TITLE	Submitted (12-APR-2000)
JOURNAL	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetradon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetradon .
COMMENT	Location/Qualifiers
FEATURES	1..973
SOURCE	/organism="Tetradon nigroviridis" /db_xref="taxon:99883" /clone="254011" /clone_1lb="G" /note="Genoscope sequence ID : COAG254AH06SP1-end : pUC-ori"
BASE COUNT	184 a 306 c 247 g 227 t 9 others
ORIGIN	
Query Match	4.9% Score 97.8; DB 17; Length 973;
Best Local Similarity	48.1% Pred. No. 4.4e-10;
Matches 337; Conservative	4; Mismatches 346; Indels 14; Gaps 2;

D	b	358	TGCTTTTGTTCCTTTATGTAGACACGTCAGAGTGTAAGTCCGCAGCGGTGGAGANGG	299
Oy	1871	TGGTACTGTGTCTCAACCTCCTGCCATCAGCCTTGCTACTTGGCCCCAAGTCTACA	1930	
D	b	238	TGGTATCATCTCTACCTCCACATCACAGGACCTCTGCGCTCATCTTTTCCAAATCTTACA	239
Oy	1931	TGATCTCTTTCACCCGAGCGCACACGCGCCGCTACTTTCACAGCATGATCCAGGGCT	1990	
D	b	238	TATATTCTTTCACAAAAGGACAAATAATACCAGAGCGCTTCAGGAAGACTGTATAGAT	179
Oy	1991	ACACCA	1996	
D	b	178	ACTCCA	173
R	E	S	U	L
RESULT 7	A1742401/c	562 bp	mRNA	linear EST 19-DEC-1999
LOCUS	A1742401			
DEFINITION	wg40e02.x1 Soares NSF_F8_9W_OT_PA_P.S1 Homo sapiens cDNA clone IMAGE:2367578.3, similar to FR:093557.093557 PUTATIVE ODORANT RECEPTOR ; mRNA sequence.			
ACCESSION	A1742401			
VERSION	A1742401.1	GI:5110689		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.			
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index			
JOURNAL	Unpublished (1997)			
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs@email.nih.gov This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Insert Length: 840 Scd Error: 0.00 Seq primer: -400P from Gldco High quality sequence stop: 468.			
FEATURES	location/Qualifiers			
source	1..562			
	/organism="Homo sapiens"			
	/db_xref="taxon:3606"			
	/clone_image="IMAGE:2367578"			
	/clone_id="Soares_NSF_F8_9W_OT_PA_P.S1"			
	/lab_host="DH10B"			
	/note="Organ: pooled; Vector: pTR3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and clonesids: Soares NbHSF pool 1: 309384-310919, 323208-325895 Soares NbZHP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-152327 Soares NbZHR8-9W pool 1: 758880-760583, 772104-774407 Soares NbHRA pool 1: 304776-306311, 320136-322833, 326280-326663 Soares NbHOT pool 1: 723720-726407, 735080-740999 Subtraction by Benito Soares and M. Fatima Bonaldo."			
BASE COUNT	116 a	161 c	167 g	117 t
ORIGIN	1 others			
Query Match	4.8%	Score 95.6	DB 9;	Length 562;
Best Local Similarity	56.8%	Pred. No. 9.9e-10;		
Matches 176;	Conservative 0;	Mismatches 134;	Indels 0;	Gaps 0;
Oy	1699	CTGGCTCAGAGGNGGGTTTCAGGCTGCCCMACATGGGCAAGAAGCTGCGCACCACTAC	1758	
D	b	439	CTCTCTTCATCAGTGCCTTGTCCGCGACGCTACCTCGGGSTTAAGGACTTCCACAGAACATAC	440

OY	1759	AACGAGGCCAAGTTCATCACCCTCACATGACTCTTTATTCCCAACCGCTCTCCCTC	1818
Db	439	AACGAGGCCAAATGTGTGCACCTTCACCGCTCTCTTAACTTGCTGTCCGATTCGGCTTC	380
OY	1819	TGCACCTTCATGTCGTGCTACACGCGGGGTGTGTGGTCACCATGTGTGAACCTTGTGTA	1878
Db	379	TTTCACCAAGGCCACGCGTCTCGAGACGGCCAAGTACCTGCTCCCTGCGGCCAACMTGTATGGCTGGG	320
OY	1879	GTCGTACAACCTCCATGGGCGCATCAGCTCTGGGGTACTTTCGGGCCCAAGTGCATATGATCTGC	1938
Db	319	CTGAGCAGGCTTGAGACAGCGGCTTTCGGTGGGTATTTTCTGTGCTTAAGTGTACGTATGATCTTC	260
OY	1939	TTCTTACCGGAGGCGCAACAGCGCCGGCTCTACTTTCACAGCATGATCCAGGGCTTACACCAATG	1998
Db	259	TGCCGCCCAACACTCAACAGACAGAGCACTTTCAGAGGCTCCATTCATCAGGACTACAGACAGAGG	200
OY	1999	AGGAGGGAGCT 2008	
Db	199	CGCTGCGGCT 190	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
BI462667				
DEFINITION		938 bp	MRNA	linear
LOCUS				
ACCESSION		603202107.F1	NIH_MGC_97	Human sapiens CDNA clone IMAGE:5268264 5'
VERSION		BI462667		
KEYWORDS		EST.		
SOURCE		human.		
ORGANISM		Homo sapiens		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
		1 (bases 1 to 938)		
		NIH-MGC http://mgc.nci.nih.gov/ .		
		National Institutes of Health, Mammalian Gene Collection (MGC)		
		Unpublished (1999)		
		Contact: Robert Strausberg, Ph.D.		
		Email: cgapbs-remail.nih.gov		
		Tissue Procurement: Miklos Palkovits, M.D., Ph.D.		
		CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shrawati		
		Toshiyuki and Piero Carninci (RIKEN)		
		CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
		DNA Sequencing by: Incyte Genomics, Inc.		
		Clone distribution: MGC clone distribution information can be		
		found through the I.M.A.G.E. Consortium/LNLN at:		
		http://image.lnl.gov		
		Plate: LRAM11676 row: m column: 01		
		High quality sequence stop: 739.		
FEATURES		Location/Qualifiers		
SOURCE		1..938		
		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/clone="IMAGE:5268264"		
		/clone_id="NIH_MGC_97"		
		/db_host="DH10B"		
		/note="Organ: testis; Vector: pBluescriptR (modified		
		pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gcgcagc		
); Oligo-dT primed using primer 5'-TTTCTTTCTTTTCTT-3',		
		size-selected for average insert size 2.2 kb and		
		normalized to R07 5. This is a primary library enriched		
		for full-length clones and constructed using the		
		Cap-trapper method (Carninci, in preparation). Library		
		constructed by M. Brownstein (NIH/NHGRI, National		
		Institutes of Health). Note: this is a NIH_MGC Library."		
BASE COUNT		186 a	260 c	294 g
ORIGIN				198 t
Query Match		4.6%	Score 92.8	DB 13
Best Local Similarity		61.7%	Pred. No. 4.8e-09	Length 938
Matches 148; Conservative		0	Mismatches 92	Indels 0
			Gaps 0	

```

/organism="Mus musculus"
/strain="NIH Swiss"
/db_xref="taxon:10090"
/clone="IMAGE:1260595"
/clone_1ib="Stratagene mouse heart (#937316)"
/sex="pooled"
/tissue_type="heart"
/dev_stage="13 day embryos"
/lab_host="SODR (Kanamycin resistant)"
/note="Organ: heart; Vector: pBluescript SK-; Site:1: EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer: Oligo dt. 93 pooled NIH/Swiss 13 day embryo hearts. Average insert size: 1.0 kb; Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGACGACG 3' -3' adaptor sequence: 5' CTCGACGTTTTTTTTTTTTTTT 3'"

```

BASE COUNT	59 a	101 c	87 g	51 t
ORIGIN				
Query Match				
Best Local Similarity	59.2%	4.1%	Score 82.8;	DB 9; Length 298;
Matches 141; Conservative	0;	0;	Pred. No. 3.6e-07;	
			Mismatches 97;	Indels 0; Gaps 0;
QY	2	TCACCTACAGGGCCATTCACGATGAGTCGCGACAGACAAGTGGCGTTCGCCGCTTTCGTCGC	61	
Db	291	TCAGCTATGAGGCGAGCGACGCGATCTCCAGTGGGAGGCGCAAGTTCCCGTCTTCTTCC	232	
QY	62	GTACACACACCGACGCGCCACACACACGTCGAGGCGCATGGTCGAGCTGATGTGCACCTTCC	121	
Db	231	GCACCATCCCGACGAGATTAAGTACACGAGTGGAAGTATAGTGGGGGTGTGTGCGACAGCTTCG	172	
QY	122	GCTGGAACCTGGATCAATTGTGCTGTGAGACGAGCACCTATGCGCGCGCAATATGGCCACG	181	
Db	171	GCTGGGTCTGGATTCGCGCTCGTGTGGCACCTATGATGATACGCGGACGTCGGGCGGTACAG	112	
QY	182	TGCTTGGGCGAGCGGTGGCGCCGCGCGACGATGTGATCGCCCTTCCAGAGACGTGCGC	239	
Db	111	CGCTGGAGGAGCTGGCCCACTCCACGCGGCGCATTGTGCTGCGCTTTCACAGGACGTGTGGTC	54	

RESULT 11	
BB618551/c	
LOCUS	BB618551 666 bp mRNA linear EST 26-OCT-2001
DEFINITION	BB618551 RIKEN full-length enriched, 8 days embryo Mus musculus
	CDNA clone 5730408m21 5', mRNA sequence.

ACCESSION	BB618551	GI:16458094
VERSION	BB618551.1	
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE
AUTHORS
Arakawa, T., Cardinaci, P., Fukuda, S., Furuno, M., Hanabaki, T., Hara, A.,
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 666)

TITLE
JOURNAL
COMMENT

Hiramoto,O., Horii,F., Ishiy,I., Ito,M., Kawai,T., Konno,H., Kouda,M., Koyota,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasak,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeeda,Y., Tanaka,T., Toyota,I., Mutamatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESIS (Atakawa,T., et al. 2001)
Unpublished (2001)
Contact: yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gscc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.,
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.*, 10 (10), 1617-1630 (2000)
wagii, K., Fujiwara, S., Inoue, K., Togawa, Y., Itawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.,
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.,
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamanaoka, I., Altzawa

K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y. Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. *Nature. Genome.* 12, 673-677 (2001). Please visit our web site (<http://genome.igsc.riken.go.jp>) for further details.

FEATURES

source Location/Qualifiers

1. 666

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="5730408M21"
/clone_1lb="RIKEN full-length enriched, 8 days embryo"
/sex="mixed"
/dev_stage="8 days embryo"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory I, RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA w. primed with a primer [5' GAGGAGAGAGAGATCCAGAGGCTCTTTTTTTTTTTTTTAA 3'], cDNA w. prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length transcripts. cDNA went through one round of subtraction. Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGGAGAGATTCGAGATTATTAATTAATATCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified plasmidscript KS(+) after bulk excision from Lambda I. Cloning sites, 5' end: SalI; 3' end: BamHI."

```

BASE COUNT	152 a	189 c	181 g	144 t	ORIGIN
Query Match	4.1%;	Score 82.2;	DB 10;	Length 666;	
Best Local Similarity	54.0%;	Pred. No. 6.8e-07;			
Matches 168;	Conservative	0;	Mismatches 143;	Indels	0; Gaps
QY 1697	TGCGGCTCAGGAGGGGTGCTTACGCTTGGGCTACAGGCGAAGAGAGCTCCGACCAACT	1756			
Db 571	TTCTCTCTCATACACACACTTTGCTGTGCTAGACTCTGGTAAAGAACTCGGGAGAACT	512			
QY 1757	ACAAGAGGCGAAGTTCATCACTCCATGACCTTCTATTTCACCTCATCCGCTCC	1816			
Db 511	ATAAGAGGCAAAATGTGTCACCTTACAGCTGTGCTCTCACTTGATGCTGATGCTT	452			
QY 1817	TCTGCACCTTCATGTGTGCTTACAGCGGGGTGCTGCTACCATGTGAGACCTTTGGTCA	1876			
Db 451	TCTTCACACATGTCAAGACATTTACAGGGCGAGCTACCTACCGCGGTCAATGTGCTGGAG	392			
QY 1877	CTGTGCTCAACCTCCCTGGGCGATCAGACCTGGGCTACTTGGGCCCAACTGTATAGATATCC	1936			
Db 391	GGCTGGCACTTGAGATGGCGGCTTACAGCGGTATTTCTCCCTAAATGGTATGATATTC	332			
QY 1937	TCTTTACCGCGAGGCGACACAGCGCGGCTACTTCAACAGATGATCCAGGCTACAGCA	1996			
Db 331	TCTGCCGTCCAGACTCAACACACAGAACACTTTCAGGCTTCATCTCAAGACTTACAGCA	272			
QY 1997	TGAGGAGGGGAC 2007				
Db 271	GGGCGTGGCGC 261				
RESULT 12					
AZ535744					
LOCUS	AZ535744	816 bp	DNA	linear	GSS 03-NOV-2009
DEFINITION	ENTMO25R Entomobea histolytica Sheared DNA Entomobea histolytica				
ACCESSION	AZ535744				
VERSION	AZ535744.1	GI:11092691			

KEYWORDS	GSS.
SOURCE	Entamoeba histolytica.
ORGANISM	Entamoeba histolytica
REFERENCE	Eukaryota; Entamoebidae; Entamoeba.
AUTHORS	1 (bases 1 to 816)
TITLE	Loftus,B., Van Aken,S. and Fraser,C. Determination of clone end sequences from Entamoeba histolytica
JOURNAL	HMI:IMSS sheared DNA library
COMMENT	Unpublished (2000) Contact: Brendan J Loftus Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0208 Fax: 301 838 3543 Email: bjoftus@tigr.org Clones are derived from the Entamoeba histolytica HMI:IMSS sheared DNA library
FEATURES	Seq primer: M13-Reverse Class: Shotgun High quality sequence start: 36 High quality sequence stop: 816. Location/Qualifiers 1..816 /organism="Entamoeba histolytica" /strain="HMI:IMSS" /db_xref="taxon:5759" /clone_lib="Entamoeba histolytica Sheared DNA" /note="Vector: pHOSt1. Site.1: Bst I; Constructed at The Institute for Genomic Research (TIGR), Rockville, MD. Genomic DNA isolated from broth cultures of E. histolytica using a method described by Clark and Diamond (Clark, C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a method for isolate identification. Exp. Parasitol. 77:450.). The DNA was mechanically sheared to give a tight size distribution (~2 kb). The v + i method used for the library construction is described in detail in Smith, H.O. and Venter, J.C. (Making small insert libraries for whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barell, Oxford University Press, 1999)."
BASE COUNT	125 a 265 c 54 g 372 t
ORIGIN	
Query Match	4.0%; Score 80.8; DB 17; Length 816;
Best Local Similarity	45.5%; Pred. No.1.5e-06;
Matches 286; Conservative 0; Mismatches 342; Indels 0; Gaps 0;	
QY 1372	GTCCTCACCTGCTTGGCCGAGGCCCTCTTCCCTGCTGCTTCACAATTGCACTGCC 1431
DB 100	GTCCTCACCTGCTTGGCCGAGGCCCTCTTCCCTGCTGCTTCACAATTGCACTGCC 159
QY 1432	TGTATCGCGCGGCTTCTTCAGATCGCTGCGCGCTTCAGATGAGGCGCGCTTCCCA 1491
DB 160	TTCCTTCATCATCATCTTTCATCATCTTTCATCATCATCTTTCATCATCATC 219
QY 1492	CAGCGCTTACAGCTACTGGGTCGGCTACCGAGGGCCCTACGTCCTTANGGCAATTATCAG 1551
DB 220	GTCCTTCATCATCATCTTTCATCATCTTTCATCATCATCTTTCATCATCATC 279
QY 1552	GTATCCAAATAGTCATGTGTGTATGTGGANMGCTGGAGGAGGCCCTCACTCCACCCCGT 1611
DB 280	GTCCTTCATCATCATCTTTCATCATCTTTCATCATCTTTCATCATCATCTTTCATCATC 339
QY 1612	ACTGACCCCGATGAGCCCAAGATCACAATTGTCTTCCGTAAACCCCACTACCGCAACAGC 1671
DB 340	GTCCTTCATCATCATCTTTCATCATCTTTCATCATCTTTCATCATCATCTTTCATCATC 399
QY 1672	CTGCTGCTTACACACAGCGCTGAGCTGCTGCTCTCAGTGGTGGTTTACGTTGCGCTTAC 1731
DB 400	GTCCTTCATCATCATCTTTCATCATCTTTCATCATCTTTCATCATCATCTTTCATCATC 459
QY 1732	ATGGGCAAGAGCTGGCCACCAACATACAGAGGCCCAAGTTCATCATCATCATCATCATGAC 1791

D _b	460	TTCTAATCAGCGCTTCTTCATCAGTCGTTCTTCACAGCGTCTTCTTCATCGTCTTCTTCATCGTC	519
Q _Y	1792	TTCTATTTCACATTCATCCGCTCTCCCTCTGCACCTTATGTCTGACAGCGGGGTGCTG	1851
D _b	520	TTCTTCATCGCTTCTTCACAGCGTCTTCTTCATCAGTCTTCTTCATCGTCTTCTTCACGCTC	579
Q _Y	1852	GTCCACATCGGTGAGACCTCTTGGTCACGTGCTCAACCTCTGGACATAGCCTGGGGCTAC	1911
D _b	580	ATCATCATCTCTTCTTCATCATCATCGTCACTCTTCATCATCATGTGATCTTCATTCATCATC	639
Q _Y	1912	TTTGGCCCCAAGTCCATCATGATTCGCTCTTCTTACCCGGAGCGGACACGCGCCGCACTATC	1971
D _b	640	TTCTTCATCATCATCTTCTTCATCATCATCACTCTTCTTATCATCATCTTCTTCATCATCATC	699
Q _Y	1972	AACGACATGATCCAGGGCTACACCATGTA	1999
D _b	700	GTCACTTCATCATCATCATCTTCTTCATCA	727

RESULT 13	AA853967/c	496 bp	mRNA	linear	EST 31-DEC-1998
LOCUS	AA853967				
DEFINITION	AA853967				
ACCESSION	AA853967				
VERSION	AA853967				
KEYWORDS	AA853967.1	GI:2941505			
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 496)				
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.				
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index				
COMMENT	Unpublished (1997)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgapsb-remail.nih.gov				
	CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo				
	, Ph.D.				
	CDNA Library Arrayed by: Greg Lennon, Ph.D.				
	DNA Sequencing by: Washington University Genome Sequencing Center				
	Clone distribution: NCI-CGAP clone distribution information can be				
	found through the I.M.A.G.E. Consortium/ILNL at:				
	www.bio.llnl.gov/dbnp/image/image.html				
FEATURES					
SOURCE					
	Possible reversed clone: similarity on wrong strand				
	Insert Length: 878 Std Error: 0.00				
	Seq primer: -40ml3 fwd. ET from Amersham				
	High quality sequence stop: 223.				
	Location/Qualifiers				
	1..496				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone_image="1393866"				
	/clone_id="Soares_testis_NHT"				
	/sex="male"				
	/lab_host="DH10B"				
	/note="Vector: pUT73D-Pac (Pharmacia) with a modified				
	polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA				
	was prepared from mRNA obtained from Clontech Laboratories				
	, Inc. and primed with a Not I - oligo(dT) primer [5'				
	TGTTCACATCTGCAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].				
	Double-stranded cDNA was ligated to Eco RI adaptors				
	(Pharmacia), digested with Not I and cloned into the Not I				
	and Eco RI sites of the modified pUT73 vector. Library				
	went through one round of normalization to Cots, and was				
	constructed by Bento Soares and M. Fatima Bonaldo."				
BASE COUNT	102 a	145 c	148 g	100 t	1 Others
ORIGIN					

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 577)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamll,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE

Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts

JOURNAL

Unpublished (2000)

COMMENT

Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunne@genetics.utah.edu

Insert length: 10000 Std Error: 0.00

Plate: 0076 row: F column: 07

Seq primer: CACACAGCAACACGCTATGACC

Class: plasmid ends

High quality sequence stop: 577.

FEATURES

Source

1..577
Location/Qualifiers

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUC1M0076F07"

/clone_1lb="Mouse 10kb plasmid UUC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PMD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of PMD42 (g1473214|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT

144 a 142 c 178 g 113 t

ORIGIN

Query Match

3.8%; Score 76.4; DB 17; Length 577;

Best Local Similarity 75.4%; Pred. No. 1e-05;

Matches 95; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 981 GATCCCTATGTCATGTCCTCAAGAGTCCAGTCCAGGCAAAAGAGAGCTGTGGG 1040
DB 472 GTTCCCATATCCATGTCCTCAAGAGTCCAGTCCAGGCAAAAGAGAGCTGTGGG 413
QY 1041 CATCCAGTCTGCTGCTGAGTGCATGATGCTCCGCGGCACTTCCATCAACCAAC 1100
DB 412 CCTCCACCCGCTGCTGAGTGTGAGCTGTCGCGGCACTTCAACCAACGATC 353
QY 1101 TGAATG 1106
DB 352 AGTAGG 347

Search completed: May 23, 2003, 18:10:19

Job time : 2843.62 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.P5_A578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 12:51:29 ; Search time 69.9475 Seconds

(Without alignments)

8812.617 Million cell updates/sec

Title: US-09-927-315-12

Perfect score: 2010

Sequence: 1 atcacatcagcgccatcag.....acacatgagagggactag 2010

Scoring table:

IDENTITY_NUC

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCtUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/Backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2010	100.0	2010	4	US-09-361-631-8 Sequence 8, Appl1
2	1232.8	61.3	2993	4	US-09-361-631-3 Sequence 3, Appl1
3	1212.6	60.3	2532	4	US-09-361-631-4 Sequence 4, Appl1
4	187	9.3	4134	4	US-09-162-021B-1 Sequence 17, Appl1
5	187	9.3	4134	4	US-09-687-477-17 Sequence 1, Appl1
6	172.6	8.6	5275	1	US-08-485-588-1 Sequence 1, Appl1
7	172.6	8.6	5275	1	US-08-485-588-1 Sequence 1, Appl1
8	172.6	8.6	5275	2	US-08-480-751-1 Sequence 1, Appl1
9	172.6	8.6	5275	2	US-08-943-986-1 Sequence 1, Appl1
10	172.6	8.6	5275	3	US-08-353-784-1 Sequence 1, Appl1
11	172.6	8.6	5275	4	US-08-484-719B-1 Sequence 1, Appl1
12	172.6	8.6	5275	4	US-08-484-719B-1 Sequence 1, Appl1
13	167	8.3	3809	1	US-08-485-588-3 Sequence 3, Appl1
14	167	8.3	3809	1	US-08-484-751-3 Sequence 3, Appl1
15	167	8.3	3809	2	US-08-480-751-3 Sequence 3, Appl1
16	167	8.3	3809	2	US-08-943-986-3 Sequence 3, Appl1
17	167	8.3	3809	3	US-08-353-784-3 Sequence 3, Appl1
18	167	8.3	3809	3	US-08-484-719B-3 Sequence 3, Appl1
19	167	8.3	3809	4	US-08-546-998-2 Sequence 3, Appl1
20	167	8.3	3809	4	US-08-484-159-3 Sequence 3, Appl1
21	167	8.3	5006	1	US-08-485-588-2 Sequence 2, Appl1
22	167	8.3	5006	1	US-08-484-565-2 Sequence 2, Appl1
23	167	8.3	5006	2	US-08-480-751-2 Sequence 2, Appl1
24	167	8.3	5006	2	US-08-943-986-2 Sequence 2, Appl1
25	167	8.3	5006	3	US-08-353-784-2 Sequence 2, Appl1
26	167	8.3	5006	3	US-08-484-719B-2 Sequence 2, Appl1
27	167	8.3	5006	4	US-08-546-998-1 Sequence 1, Appl1

28	167	8.3	5006	4	US-08-484-159-2 Sequence 2, Appl1
29	146.8	7.3	3384	2	US-08-687-289A-1 Sequence 1, Appl1
30	139.2	6.9	4131	1	US-08-485-588-4 Sequence 4, Appl1
31	139.2	6.9	4131	1	US-08-484-565-4 Sequence 4, Appl1
32	139.2	6.9	4131	2	US-08-480-751-4 Sequence 4, Appl1
33	139.2	6.9	4131	2	US-08-943-986-4 Sequence 4, Appl1
34	139.2	6.9	4131	3	US-08-353-784-4 Sequence 4, Appl1
35	139.2	6.9	4131	3	US-08-484-719B-4 Sequence 4, Appl1
36	139.2	6.9	4131	4	US-08-484-159-4 Sequence 4, Appl1
37	124.2	6.2	3177	4	US-09-134-513-1 Sequence 13, Appl1
38	96	4.8	2595	4	US-09-619-353-13 Sequence 2, Appl1
39	84.6	4.2	4000	2	US-08-687-289A-2 Sequence 2, Appl1
40	81.8	4.1	2877	4	US-09-619-353-1 Sequence 1, Appl1
41	71.6	3.6	2148	5	PCt-US93-01642-1 Sequence 1, Appl1
42	64	3.2	2426	5	PCt-US91-09422-20 Sequence 20, Appl1
43	58.2	2.9	4095	5	PCt-US91-09422-18 Sequence 18, Appl1
44	58.2	2.9	2619	3	US-08-337-797A-1 Sequence 1, Appl1
45	58.2	2.9	2619	3	US-08-337-797A-3 Sequence 3, Appl1

ALIGNMENTS

```
RESULT 1
US-09-361-631-8
; Sequence 8, Application US/09361631
; Patent No. 6383778
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeyer, Juergen
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-protein Coupled Receptor
; FILE REFERENCE: 02307E-08872005
; CURRENT APPLICATION NUMBER: US/09361,631
; EARLIER FILING DATE: 1999-07-27
; EARLIER APPLICATION NUMBER: US 60/095,464
; EARLIER FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: US 60/112,747
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 8
; LENGTH: 2010
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) B4
US-09-361-631-8
; OTHER INFORMATION: nucleotide sequence

Query Match      100.0%; Score 2010; DB 4; Length 2010;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATCACTAGAGGCGCATGCGGATGAGTCGAGCAAGAGTGGCTTCCGGCTTGGC 60
      |||
DB      1 ATCACTAGAGGCGCATGCGGATGAGTCGAGCAAGAGTGGCTTCCGGCTTGGC 60
      |||
QY      61 CGTACACACCCAGCGCCAGCACACGTCGAGGCGCATGAGTGTGATGTCACCTTC 120
      |||
DB      61 CGTACACACCCAGCGCCAGCACACGTCGAGGCGCATGAGTGTGATGTCACCTTC 120
      |||
QY      121 CGCTGAGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 180
      |||
DB      121 CGCTGAGTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 180
      |||
QY      181 CGGCTGGGAGCGGGTGGCGCGCGCGCGCGCATCTGATCGCTTCAGAGAGCGCGCC 240
      |||
DB      181 CGGCTGGGAGCGGGTGGCGCGCGCGCGCGCATCTGATCGCTTCAGAGAGCGCGCC 240
      |||
QY      241 ACACTGACACCCACACCAACATGACGTGAGAGAGCGCGCGCTGCTGACATTTG 300
```

```

Db 241 ACACGTGACGCCAACACAGAACATGACGTCAGAGAGGCGCCAGCCCTGATGACATTTGTG 300
QY 301 GACAAAGCGACGACGAGACAGCGCGGTGTTGTCGTTCTGCGCCAGCCTGACCTG 360
Db 301 GACAAAGCGACGACGAGACAGCGCGGTGTTGTCGTTCTGCGCCAGCCTGACCTG 360
QY 361 TACCACTTCTTCAATGAGTGTGTCGCGCCAGAACTTCAAGCGCGCGGTGATGCGCTTC 420
Db 361 TACCACTTCTTCAATGAGTGTGTCGCGCCAGAACTTCAAGCGCGCGGTGATGCGCTTC 420
QY 421 GAGTCTGCGCGCATGACGCGCGGTGTCGACAACTTCAAGCGCGCGGTGATGCGCTTC 480
Db 421 GAGTCTGCGCGCATGACGCGCGGTGTCGACAACTTCAAGCGCGCGGTGATGCGCTTC 480
QY 481 TTCTGCGCATCACCATGACGAGGTCGCGCGGTGTCGACAACTTCAAGCGCGCGGTGATG 540
Db 481 TTCTGCGCATCACCATGACGAGGTCGCGCGGTGTCGACAACTTCAAGCGCGCGGTGATG 540
QY 541 GGCACACAGGCTGGGCGCGCCACCCCTGACGACGACGACGACGACGACGACGACGACG 600
Db 541 GGCACACAGGCTGGGCGCGCCACCCCTGACGACGACGACGACGACGACGACGACGACG 600
QY 601 GAGTCCGACAACTGCTGACGACGACGACGACGACGACGACGACGACGACGACGACG 660
Db 601 GAGTCCGACAACTGCTGACGACGACGACGACGACGACGACGACGACGACGACGACG 660
QY 661 GAGGCTGCTGCTGACGAGGTCGACGAGGTCGACGAGGTCGACGAGGTCGACGAGGTC 720
Db 661 GAGGCTGCTGCTGACGAGGTCGACGAGGTCGACGAGGTCGACGAGGTCGACGAGGTC 720
QY 721 CTCCTGCGCTGTGACAAAGACGACGACGACGACGACGACGACGACGACGACGACG 780
Db 721 CTCCTGCGCTGTGACAAAGACGACGACGACGACGACGACGACGACGACGACGACG 780
QY 781 GAGAGATCTGGAAGGTCACTTCACTTCTGACACCAAACTTCTTTCGACCCGCAA 840
Db 781 GAGAGATCTGGAAGGTCACTTCACTTCTGACACCAAACTTCTTTCGACCCGCAA 840
QY 841 GGGAGCGGCTGCTGACGAGGTCGACGAGGTCGACGAGGTCGACGAGGTCGACGAGGTC 900
Db 841 GGGAGCGGCTGCTGACGAGGTCGACGAGGTCGACGAGGTCGACGAGGTCGACGAGGTC 900
QY 901 TTCGAGAGCGTGGCTCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 960
Db 901 TTCGAGAGCGTGGCTCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 960
QY 961 CTCGACACCGTCAACACACGATCCCTATGTCATGTCATGTCATGTCATGTCATGTCAT 1020
Db 961 CTCGACACCGTCAACACACGATCCCTATGTCATGTCATGTCATGTCATGTCATGTCAT 1020
QY 1021 CAAAAGAGAGAGCTGAGGATCAGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1080
Db 1021 CAAAAGAGAGAGCTGAGGATCAGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1080
QY 1081 GGCACCTTCTTCAACACACGATCCCTATGTCATGTCATGTCATGTCATGTCATGTCAT 1140
Db 1081 GGCACCTTCTTCAACACACGATCCCTATGTCATGTCATGTCATGTCATGTCATGTCAT 1140
QY 1141 TCTGCTTCAAGCGGACGCTGTCTTCTGTAATGTCATGTCATGTCATGTCATGTCATG 1200
Db 1141 TCTGCTTCAAGCGGACGCTGTCTTCTGTAATGTCATGTCATGTCATGTCATGTCATG 1200
QY 1201 GGCCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
Db 1201 GGCCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1260
QY 1261 CACTTCCAGACACCATAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db 1261 CACTTCCAGACACCATAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
QY 1321 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
Db 1321 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380

Db 1321 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
QY 1381 TGCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db 1381 TGCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
QY 1441 GTGCTTCTTCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
Db 1441 GTGCTTCTTCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
QY 1501 AGTACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1560
Db 1501 AGTACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1560
QY 1561 ATGCTATTTGTTGTAATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 1620
Db 1561 ATGCTATTTGTTGTAATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 1620
QY 1621 GATGACCCCAAGATCACAATTTGCTCTGTAACCCCAACTACCGCAACAGCCTGCTTC 1680
Db 1621 GATGACCCCAAGATCACAATTTGCTCTGTAACCCCAACTACCGCAACAGCCTGCTTC 1680
QY 1681 AACACAGCGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Db 1681 AACACAGCGCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
QY 1741 GAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1800
Db 1741 GAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1800
QY 1801 ACCTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
Db 1801 ACCTATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
QY 1861 GTGACCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
Db 1861 GTGACCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
QY 1921 AAGTCTACATGATCTCTTCTTACCGGAGGCGCAACGCGCGCGCTTCAACAGCATG 1980
Db 1921 AAGTCTACATGATCTCTTCTTACCGGAGGCGCAACGCGCGCGCTTCAACAGCATG 1980
QY 1981 ATCCAGGCTACACCATGAGGAGGACTAG 2010
Db 1981 ATCCAGGCTACACCATGAGGAGGACTAG 2010

RESULT 2
US-09-361-631-3
: Sequence 3, Application US/09361631
: Patent No. 6383778
: GENERAL INFORMATION:
: APPLICANT: Zuker, Charles S.
: APPLICANT: Adler, Jon Elliot
: APPLICANT: Lindemeyer, Juergen
: APPLICANT: The Regents of the University of California
: TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
: TITLE OF INVENTION: Involved in Sensory Transduction
: FILE REFERENCE: 02307E-088720US
: CURRENT APPLICATION NUMBER: US/09/361,631
: EARLIER FILING DATE: 1999-07-27
: EARLIER APPLICATION NUMBER: US 60/095,464
: EARLIER FILING DATE: 1998-07-28
: EARLIER APPLICATION NUMBER: US 60/112,747
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 3
: LENGTH: 2993
: TYPE: DNA
: ORGANISM: Rattus sp.
: FEATURE:
: OTHER INFORMATION: rat G-protein coupled receptor (GPCR) B4
```

OTHER INFORMATION: nucleotide sequence
US-09-361-631-3

Query Match 61.3%; Score 1232.8; DB 4; Length 2993;
Best Local Similarity 76.9%; Pred. No. 1.9e-255;
Matches 1568; Conservative 0; Mismatches 442; Indels 30; Gaps 4;

1 ATACCTACAGGCGCCATACAGGATGAGCTGCGAGACAGAGTGGCTTCCCGCTTGGCTG 60
Db 505 ATACATATACAGCGCATCTCCGAAAGCTGGGACACAGGCGACATCCCTACATGCTA 564
QY 61 CGTACACACACCGCCGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 120
Db 565 CGCAGAGTGGCCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 624
QY 121 CGCTGACATCGATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db 625 CAATGGAACCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 684
QY 181 CTGCTTGGCGAGCGCTGGCC---CGGCGGACATCTGATCGCTTCCAGAGAGCGTG 237
Db 685 CTGTGAGCGAGCGCTGACACAAAGAGCGACATCTGCTTCCAGAGAGCTTCTG 744
QY 238 CCCACATCGAGCCACACGACATGACGTGACAGAGAGCGACGCGCTGCTGACCAT 297
Db 745 CCCATCTGAGTCCAGCGAGGTGATGAGTCCGAGAGAGACGACATGACGACATC 804
QY 298 GTGGAACAGCTGACAGAGACAGAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 357
Db 805 CTGGAACAGCTGCGGCGAGCTGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 864
QY 358 CTGTACCATCTTCTCATAGAGTGTGCTGCGACAGACTTACAGGCGCGCGCTGCTGCT 417
Db 865 CTGTATAGCTTCTTTCACAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 924
QY 418 TCCGAGTCTGCGCATCGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 477
Db 925 TCTGAGTCTGCGCTATCGACCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 984
QY 478 ACCTTCTGGGCGATCAACATCCAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 537
Db 985 ACTTCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1044
QY 538 TGGGCGCCACAGGCTGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 597
Db 1045 CGCCGTAACAGGCGAGGATCCCGTGAACAGAGACAGCAACCTGGGAGAGACTGTAAC 1104
QY 598 CAGGAGTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
Db 1105 CAGGAGTGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1164
QY 658 GGGGAGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
Db 1165 GGGGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1224
QY 718 AGCTCTCTGGCTGTGCAAAAGACCTGACCAAGAGAGGCTGCTGCTGCTGCTGCTGCT 777
Db 1225 AGACTCTCTGGCTGTGCAAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1284
QY 778 CTGAGAGATCTGGAAGTCAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
Db 1285 CTGAGAGATCTGGAAGTCAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1344
QY 838 CAGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
Db 1345 CAGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1404
QY 898 CCTTCCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954
Db 1405 CCTTCCAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1464
QY 955 ACATCTCTGACACCGCTCAACACAGATCCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014

Db 1465 GTGTCTGTGTACACCCCAACACAGGCTCCGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1524
QY 1015 TCAGGGCAAAAGAGAGCGCTGTGGCATCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 1074
Db 1525 CCAGGGCAAAAGAGAGCGCTGTGGCATCCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 1584
QY 1075 CTTCGGGCAACCTCTCTCAACCACTGAA-----TGCCCGAAT 1113
Db 1585 ATGCAAGGACACTCTCAACCGCTGACAGATGATTAAGTCTGCTGCTGCTGCTGCTGCT 1644
QY 1114 AAGAGTGTCTCTACAGAGTGAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1173
Db 1645 TCCATGTGTCTCTACAGAGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1704
QY 1174 TGGCAGAGGACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1233
Db 1705 TGGCAGAGAGTGGCCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1764
QY 1234 CTGGCCATCTGTGATATCTGAGAGCACTTCCAGACACCACTGCTGCTGCTGCTGCTGCT 1293
Db 1765 CTGGCCATCTTCTTCACTCTTGTGAGACATTTCCAGACACCACTGCTGCTGCTGCTGCT 1824
QY 1294 GGGCCATGTGCTTCTGATGCTGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1353
Db 1825 GGGCCATGTGCTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1884
QY 1354 TACGTGGGCGCGCCAGGCTTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1413
Db 1885 TATGTGGGCGCGCCAGGCTTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1944
QY 1414 TTCACAATTTGCACTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1473
Db 1945 TTCTCATCTGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2004
QY 1474 ATGGCCAGCGCTTCCAGCGCGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1533
Db 2005 ATGGCCAGCGCGCTTCCAGCGCGCTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2064
QY 1534 TGTATGGCAATTAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1593
Db 2065 TTGTGCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2124
QY 1594 CCTCACTCCACCC---CCGTATGACCCCGATGACCCCAAGATCAATTTGCTGCTGCT 1650
Db 2125 ACATCAACCCATTTGGCGGAGCCGAGATGACCCCAACATATGATTCCTGCTGCTGCT 2184
QY 1651 AACCCCACTACCGCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1710
Db 2185 CACCCCTAATACCGCAACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2244
QY 1711 GTGGGTTTACGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1770
Db 2245 CTGGGTTTACGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2304
QY 1771 TTTCATACCCCTGACATGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1830
Db 2305 TTTCATACCTGACATGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2364
QY 1831 TCTGCTTACAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1890
Db 2365 TCTGCTTACAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2424
QY 1891 CTGGCCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1950
Db 2425 CTGGCCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2484
QY 1951 CGCAACAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2010
Db 2485 CGCAACAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2544

RESULT 3
US-09-361-631-4

```

; Sequence 4, Application US/09361631
; Patent No. 6383778
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeier, Juergen
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
; TITLE OF INVENTION: Involved in Sensory Transduction
; FILE REFERENCE: 02307E-088720US
; CURRENT APPLICATION NUMBER: US/09/361,631
; CURRENT FILING DATE: 1999-07-27
; EARLIER APPLICATION NUMBER: US 60/095,464
; EARLIER FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: US 60/112,747
; EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentl Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2532
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) B4
; OTHER INFORMATION: nucleotide sequence
US-09-361-631-4

Query Match      60.3%; Score 1212.6; DB 4; Length 2532;
Best Local Similarity 76.3%; Pred. No. 3.9e-251;
Matches 1555; Conservative 0; Mismatches 454; Indels 30; Gaps 4;

QY 2 TCACCTACAGCGCCATGACGATGAGCTGCGAGACAAAGGTGCGCTTCCCGGCTTGTCTGC 61
DB 494 TCACATATAGCGCCATGACGATGAGCTGCAAGAACAGCGCGCTTCCCTGCATGTCTGC 553
QY 62 GTACCAACCCAGCGCCAGCCAGCATGCGAGCGCATGTGTGCTGATGTCTGCACTTCC 121
DB 554 GCACGTGCGCCAGCGCCAGCCAGCATGCGAGCGCATGTGTGCTGATGTCTGCACTTCC 613
QY 122 GCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 181
DB 614 AGTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 673
QY 182 TGTCTGGGAGCGCGGTGC---CCGGCGCAGCATCTGCACTGCGCTTCCAGAGAGAGCTGC 238
DB 674 TGTGAGCGAGCGCTGCAACACATGCGCATCTGCACTGCGCTTCCAGAGAGAGCTGC 733
QY 239 CCACACTGACAGCCCAACCAACATGATGATGATGATGATGATGATGATGATGATGATG 298
DB 734 CCGTACCGAAGCCCAACCAACATGATGATGATGATGATGATGATGATGATGATGATG 793
QY 299 TGGACAAGCTGACAGACAGACAGCGCGCTGATGATGATGATGATGATGATGATGATGATG 358
DB 794 TGGACAAGCTGCGCGGAGCTTGGCGCGCTGATGATGATGATGATGATGATGATGATGATG 853
QY 359 TGTACCACTTCTTAATGAGTGTGCGCGCAAACTTCAAGCGCGCGCTGTGATGCGCT 418
DB 854 TGCACCACTTCTTAATGAGTGTGCGCGCAAACTTCAAGCGCGCGCTGTGATGCGCT 913
QY 419 CCGAGTCTGGGCGCATGAGCGCGCTGCGCAACCTGAGGAGCGTGGGCGCACTTGGGCA 478
DB 914 CTGAGTCTGGGCGCATGAGCGCGCTGCGCAACCTGAGGAGCGTGGGCGCACTTGGGCA 973
QY 479 CTTTCTTGGGCGCATGAGCGCGCTGCGCAACCTGAGGAGCGTGGGCGCACTTGGGCA 538
DB 974 CTTTCTTGGGCGCATGAGCGCGCTGCGCAACCTGAGGAGCGTGGGCGCACTTGGGCA 1033
QY 539 GGGGCGCCAGAGCGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 598
DB 1034 GCCATGACAAACCCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1093
QY 599 AGGAGTGGAGAACTGCTGTAAGCGCACTGTCTTCAACCACTTGTGAGGCTGTG 658
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 1094 AGGATCGAGCGCTGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1153
QY 659 GGGAGCGGTGCTTACAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 718
DB 1154 GGGAGCGGTGCTTACAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1213
QY 719 GCTTCTGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 778
DB 1214 GATCTCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1273
QY 779 TTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 838
DB 1274 TCAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1333
QY 839 AAGGGAGCGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 898
DB 1334 AAGGGAGCGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1393
QY 899 CTTTCCAGAGCGGTGCTTACAGCGGTGATGATGATGATGATGATGATGATGATGATGATG 955
DB 1394 CTTTCCAGAGCGGTGCTTACAGCGGTGATGATGATGATGATGATGATGATGATGATGATG 1453
QY 956 CATCTCTGACAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1015
DB 1454 TGTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1513
QY 1016 CAGGGCAAAAGAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1075
DB 1514 CTGGGCAAAAGAGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1573
QY 1076 TTCCCGGACCTTCTCAACCACTGAA-----TGCCGAAATA 1114
DB 1574 CGCGGACACCTTCAACCACTGAA-----TGCCGAAATA 1174
QY 1115 ACGAGTGTCTTACAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1174
DB 1634 CAGTGTGTCTTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1693
QY 1175 GGCATGAGAGCGCCAGCATGCTGAGCGCGCTGAGCGCGCTGAGCGCGCTGAGCGCGCT 1234
DB 1694 GGCATGAGAGCGCCAGCATGCTGAGCGCGCTGAGCGCGCTGAGCGCGCTGAGCGCGCT 1753
QY 1235 TGGCATCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1294
DB 1754 TGGCATCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1813
QY 1295 GCGCCATGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1354
DB 1814 GCGCCATGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1873
QY 1355 ACGTGGGCGCGCCAGAGTGTCTCACTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCT 1414
DB 1874 ATGTGGGCGCGCCAGAGTGTCTCACTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCT 1933
QY 1415 TCACAAATTTGATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1474
DB 1934 TCTCCGCTGTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1993
QY 1475 TGGCAGCGCGGTGCGCAAGCGCGCTGAGCGCTGCTTCAAGATGATGATGATGATGATGATG 1534
DB 1994 TGGCAGCGCGGTGCGCAAGCGCGCTGAGCGCTGCTTCAAGATGATGATGATGATGATGATG 2033
QY 1535 CTATGCAATTTATCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1594
DB 2054 TGTGTGCTTATCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2113
QY 1595 CTGAGTCCAGCC---CCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1651
DB 2114 CCATCAACCGCATTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2173
QY 1652 ACCCACTACCGCAAGCGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1711
DB 2174 ACCCACTACCGCAAGCGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2233

```


	Query Match	5.3%	Score 187	DB 4:	Length 4134	
	Best Local Similarity	51.2%	Pred. No. 2,7e-31			
	Matches 464	Conservative 0	Mismatches 440	Indels 3	Gaps 1	
QY	1099	ACTGATACCCGGAATACGAGTGTCTACCAAGATGAGACCTCTGCTTCAAGGGCAG	1158			
Db	2209	ACAAGTCCCGGAATGATTTCTGTGATGATGAACACACACTCGTGCATCCCAAGGAG	2268			
QY	1159	CTGGCTTCCTCGGATATGATGATGAGCACACCACATCGGTGTGGCCCTGTGGCCGGCCG	1218			
Db	2259	ATCGATGACTCTGTCTGGACGAGAGCCCTTCCGGATTCGCTTACCATCTTCCGCCGTACTG	2328			
QY	1219	GGCCTCTCAGACACCTGGCCATCTCTGTGATATTTCTGGAGGCACTTCCAGACACCATA	1278			
Db	2339	GGCACTCTGATCACCTCTCTCTGCTGGGGGCTTTCATCAATCAAGTGAACACTCCACAC	2388			
QY	1279	GTTCTCTGGGCTGGGGGGCCCATGTGCTTCTCGATGCTGACATCGTGTGGTGGGATAC	1338			
Db	2389	GTTGAAGGCCACACAAACCGGAGATGTCTTCACTGCTGCTTCTTCCCTCATCTGCTGCTTC	2448			
QY	1339	ATGTGTGTTCGCGGTACTGTGGGGCCGCCCAAGTCTTCCACTGCTCTGCGCCAGGCC	1398			
Db	2449	TCCAGCTGCTCATTTATTCGCGAGGCCCAAGGACTGGACCTGTGCGCTTCGCCCAACCG	2508			

Query Match	9.3%	Score 187	DB 4	length 4134
Best Local Similarity	51.2%	Pred. No. 2.7e-31		
Matches 464; Conservative	0	Mismatches 440	Indels 3	Gaps 1

```

QY 1099 ACTGATGCCCCGGAATACGAGTGTGCTTACAGAGTGAGACCTCTGCTTCAAGCGGAG 1158
Db 2209 ACAAGATGCCCCGGAATGATTTGCTGCAATGAGAACACAGCTGTGATCGCCAAAGAG 2268
QY 1159 CTGCTCTCCGGAATGAGGCAATGAGGCAACCACTGCTGTGCGCTCTGCTGCGCCCTG 1218
Db 2289 ATCGAGTACCTGTGCTGAGGAGGAGCCCTGCGGAGGCTGTGACCATCTTGTGACCTATG 2328
QY 1219 GCGTTCCTCAGACACCTCGGCGCATCTGTGATATTTCTGAGAGCACTTCCAGACCCATA 1278
Db 2329 GGCATCTCGATCACCCTCTCTGTGTGAGGAGCTTTCATCAGATTGAGAAACATCTCCATC 2388
QY 1279 GTTGCCTCGCTGGGGGCCCATGCTCTCTGATGCTGACATGCTGCTGCTGCTGCTGCTG 1338
Db 2389 GTGAGAGCGCACCAACCGGAGGAGTGTCTTACCTGCTCTCTCTCTCTCTCTCTCTCTCT 2448
QY 1339 ATGAGTGTCCGGGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1398
Db 2449 TCCAGCTGCTGATCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2508
QY 1399 CTCTTTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1458
Db 2509 GCGTTTGGCATACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2568
QY 1459 GTCTGCGCTTCAAGAGTGGCCAGCGCTTCCACGCGCTTCAAGCTGCTGCTGCTGCTGCT 1518
Db 2569 CTGCTGCTTCTGCA---GGCCAGATGCTCCACAGCCCTCCACCGCAAGTGGGCGCTC 2625
QY 1519 CAGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1578
Db 2626 AACCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2685
QY 1579 GGCATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1638
Db 2686 TGCTCTGACAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2745
QY 1639 ATGCTCTCTGATACCCCACTACCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1698
Db 2746 TTCACTACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2805
QY 1699 CTGCTCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1758
Db 2806 CTCTCTGCGCGCATCTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2865
QY 1759 AACGAGGCAATTCATACCCCACTACCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1818
Db 2866 AACGAGGCAATTCATACCCCACTACCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2925
QY 1819 TGCACTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1878
Db 2926 ATCCCGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2985
QY 1879 GTGCTCAACCTGCTGCGCATACGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1938
Db 2986 CTGCGCTTCCAGCTTCTGCGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3045
QY 1939 TTCTACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1998
Db 3046 TTCAAGCGCGTGGTCAACACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3105
QY 1999 AGGAGG 2005
Db 3106 AAGGTGG 3112

```

RESULT 6
US-08-485-588-1
Sequence 1, Application US/08485588
Patent No. 5688938
GENERAL INFORMATION:

APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller

```

APPLICANT: James E. Garrett, Jr.  
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE  
TITLE OF INVENTION: MOLECULES  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: First Interstate World Center  
STREET: Suite 4700  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
COUNTRY: USA  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: FASTSEQ  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,588  
FILING DATE: 7 June, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
PRIOR APPLICATION DATA: Including application  
PRIOR APPLICATION DATA: described below: 9  
APPLICATION NUMBER: 08/353,784  
FILING DATE: 9 December, 1994  
APPLICATION NUMBER: PCT/US/94/12117  
FILING DATE: 21 October, 1994  
APPLICATION NUMBER: U.S. 08/292,827  
FILING DATE: 23 August, 1994  
APPLICATION NUMBER: U.S. 08/141,248  
FILING DATE: 22 October, 1993  
APPLICATION NUMBER: U.S. 08/009,389  
FILING DATE: 23 February, 1993  
APPLICATION NUMBER: U.S. 08/017,127  
FILING DATE: 12 February, 1993  
APPLICATION NUMBER: U.S. 07/934,161  
FILING DATE: 21 August, 1992  
APPLICATION NUMBER: U.S. 07/834,044  
FILING DATE: 11 February, 1992  
APPLICATION NUMBER: U.S. 07/749,451  
FILING DATE: 23 August, 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Hebert, Sheldon O.  
REGISTRATION NUMBER: 38,179  
REFERENCE/DOCKET NUMBER: 213/005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (213) 489-1600  
TELEFAX: (213) 955-0440  
TELEX: 67-3510  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5275 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
NAME/KEY: CDS  
LOCATION: 515..3769  
OTHER INFORMATION:  
US-08-485-588-1

```

Query Match 8.6%; Score 172.6; DB 1; Length 5275;
Best Local Similarity 50.2%; Pred. No. 3.4e-28;
Matches 455; Conservative 0; Mismatches 449; Indels 3; Gaps 1;

```

QY 1101 TGAATGCCCGGATACAGTGTGCTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1160
Db 2266 TAAATGCCCGGATACAGTGTGCTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 2325
QY 1161 GGTCTTCTGGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1220

```

```

Db 2326 CGAGTTTCGTGCGTAGACCGCCCTTGGGATGCACTACGGCTTTGGTGTGG 2385
QY 1221 CTTCCTAGACACCTGGCCATCTGCTGATATTTGAGGACATTCACACCCATAGT 1280
Db 2386 CATTTCTCTACACCTTCGCTGGGCGTCTTCATCAAGTTCCGCAACAGCCCATGCT 2445
QY 1281 TCGGTGGGCGGGGGCCCATGCTTCTGATGCTGACATGCTGCGTGGGATCAT 1340
Db 2446 CAAGGCCACCAACGGGAGCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCT 2505
QY 1341 GGTGGTCCGGTGTACGTGGGGCCGCAAGTCTCCACCTGCTCCGCGCCAGGCTT 1400
Db 2506 CAGTCCCTGTTCTTCACTGGGAGAGCCCAAGACTGACAGCTGCGCCGCGCCAG 2565
QY 1401 CTTCCTCTGCTTCAATTTGATCTCTGATGCGCGCTGCTTTTCCAGATGCT 1460
Db 2566 CTTCGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2625
QY 1461 CTGGGCTTCAAGATGGCGACCGCTTCCACGGGCTTACAGCTACTGCTGCTTACCA 1520
Db 2626 CTGGTGTGTA---GGCAAGATTCCTCCACAGCTTCCACCGGAAGTGGGGGCTTAA 2682
QY 1521 GGGGCGCTTACGCTCTCTATGCGATTTATCAGGTAATCAAAATGCTATGTAATGG 1580
Db 2683 CCTGAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2742
QY 1581 CATGCTGGACAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1640
Db 2743 GCTCAATATACAGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2802
QY 1641 TGTCCTCGTAAACCCCAACCTACGCAACAGCTGCTGCTGCTGCTGCTGCTGCT 1700
Db 2803 CATACCTGCGACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2862
QY 1701 GCTCTCAGTGTGGGTTTACGCTTCCCTACATGGGCAAGAGCTGCGCCACCATAC 1760
Db 2863 GCTGGCGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2922
QY 1761 CGAGGCCAAGTTCAACCTCAGCATGCTTCTATTTCACTGCTGCTGCTGCTGCT 1820
Db 2923 TGAAGCCAAAGTTCACTACCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2982
QY 1821 CACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1880
Db 2983 CCGCGCTTACGCACTTACGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3042
QY 1881 GCTCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1940
Db 3043 GCGCGCCAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3102
QY 1941 CTACCCGAGGCGCAACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2000
Db 3103 CAAGCCTTCCGGAACACATGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3162
QY 2001 GAGGAC 2007
Db 3163 GGTGGCC 3169

```

RESULT 7
US-08-484-565-1

Sequence 1, Application US/08484565
Patent No. 5763569

GENERAL INFORMATION:

APPLICANT: Edward M. Brown

APPLICANT: Steven E. Hebert

APPLICANT: James E. Garret, Jr.

TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE

TITLE OF INVENTION: MOLECULES

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

```

STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSBO
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,565
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5275 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 515..3769
OTHER INFORMATION:
US-08-484-565-1

Query Match 8.6%; Score 172.6; DB 1; Length 5275;
Best Local Similarity 50.2%; Pred. No. 3.4e-28;
Matches 455; Conservative 0; Mismatches 449; Indels 3; Gaps 1;

```


QY	1521	GGGGCCCTACGTCCTCATGTGGCATTTTACACGGTACACAAATGGCATTTGGTATATGG	1580
Db	2683	CTTGCAATTCCTGCTGGTCTTCTCCTGTGCACATTCATGCAGATTTGCATCTGTGCCATTTC	2742
QY	1581	CATGCTGGCAGGGCTTCATGCCACACCCCGTACTGACCCCGATGACCCCGAATATCAAT	1640
Db	2743	GCTCAATATACGGCGCCCCCTCTGAGCTACGCCAACCCAGAGCTGGAGAGACGATCATCTT	2802
QY	1641	TGTCCTCGTAACCCCACTACGCGCAACAGCCGCGTGTTCACACAGACGTAGGACCTGCT	1700
Db	2803	CATACCTCCGCACAGAGGGCTGCTCATGGCGCTGGGCTTCTGTAATGGGTACACTGCTT	2862
QY	1701	GCTCTCAGTGGTGGGTTTACGTTTCGCTCATACATGGGCAAGAGCTGCCACCACTACA	1760
Db	2863	GCTGGCGGCATCTGCTTCTTCTTCGCTTCAAGTCCCGGAAGTCCGAGAAACTTCA	2922
QY	1761	CGAGGCCAAGTATATACCCCTACGACATTCCTCATTTTACCTATCCGTCCTCCTTG	1820
Db	2923	TGAAGCCCAAGTCACTACCTTCACGATGCTCATCTTTCATATCGTGTGAATCTTTTAT	2982
QY	1821	CACCTTCATGTTCGCTTACAGCGGGGTGCTGTGCATCATGTGTGGACCTTTGGTCACTGT	1880
Db	2983	CCCCGCTTAGCGCAGCACTTACGGCAAAATTGCTGCTGCGTGGAGGTGATGGCATCT	3042
QY	1881	GCTCAACCTCTGTGGCATTCAGCCTGGGGCTACTTTCGGGCCCAAGTCTCATATGATCTT	1940
Db	3043	GCGCGCCAGCTTTGGCTGTGCTGCGCTGTATCTTCTTCAACAAGGCTCATCATCTCTT	3102
QY	1941	CTACCCGAGAGCGACACGGCCGCTCATTTCAACAGCATGATCCAGGGCTACACCATGAG	2000
Db	3103	CAACGCTTCCCGAACCATTCAGAGAGGTGGGTGCGACACACGGGGCACACGCTTAA	3162
QY	2001	GAGGGAC	2007
Db	3163	GGTGGCC	3169

RESULT 11
 US-08-484-719B-1
 Sequence 1, Application US/08484719B
 Patent No. 6031003
 GENERAL INFORMATION:
 APPLICANT: Edward F. Nemeth, Edward M.
 APPLICANT: Brown, Steven C. Hebert,
 APPLICANT: Bradford C. Van Wageningen,
 APPLICANT: Manuel F. Balandin
 APPLICANT: Forrest H. Fuller, Eric G.
 APPLICANT: Delmar, Scott T. Moe
 TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 TITLE OF INVENTION: MOLECULES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: First Interstate World Center
 STREET: Suite 4700
 STREET: 633 West Fifth Street
 CITY: Los Angeles
 STATE: California
 COUNTRY: USA
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: MS Word
 SOFTWARE: FastSeq for Windows Version 3.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,719B
 FILING DATE: 7 June, 1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/355,784
 FILING DATE: 9 December, 1994

```

1      APPLICATION NUMBER: PCT/US/94/12117
2      FILING DATE: 21 October, 1994.
3      APPLICATION NUMBER: U.S. 08/292,827
4      FILING DATE: 23 August, 1994
5      APPLICATION NUMBER: U.S. 08/141,248
6      FILING DATE: 22 October, 1993
7      APPLICATION NUMBER: U.S. 08/009,389
8      FILING DATE: 23 February, 1993
9      APPLICATION NUMBER: U.S. 08/017,127
10     FILING DATE: 12 February, 1993
11     APPLICATION NUMBER: U.S. 07/993,161
12     FILING DATE: 21 August, 1992
13     APPLICATION NUMBER: U.S. 07/834,044
14     FILING DATE: 11 February, 1992
15     APPLICATION NUMBER: U.S. 07/749,451
16     FILING DATE: 23 August, 1991
17
18     ATTORNEY/AGENT INFORMATION:
19
20     NAME: Douglas C. Murdock
21     REGISTRATION NUMBER: 37,549
22     REFERENCE/DOCKET NUMBER: 213/007
23
24     TELECOMMUNICATION INFORMATION:
25
26     TELEPHONE: (213) 489-1600
27     TELEFAX: (213) 955-0440
28     TELEX: 67-3510
29
30     INFORMATION FOR SEQ ID NO: 1:
31
32     SEQUENCE CHARACTERISTICS:
33
34     LENGTH: 5275 base pairs
35     TYPE: nucleic acid
36     STRANDEDNESS: single
37     TOPOLOGY: linear
38
39     MOLECULE TYPE: cDNA to mRNA
40
41     FEATURE:
42
43     NAME/KEY: CDS
44     LOCATION: 515..3769
45
46     US-08-484-119B-1

```

Query Match	Similarity	8.6%	Score 172.6	DB 3	Length 5275
Best Local	Similarity	50.2%	Pred. No. 3.4e-28		
Matches	455	Conservative	0	Mismatches 449	Indels 3
					Gaps 1
QY	1101	TCGATGCCCGGATTAACGAGTGGTCTCTACCAAGTAGAGACCTTCGCTTTCACAGCGCAGCT	1160		
Db	2266	TAAAGTCCCTGATGACTTCTGTGTCGAATGAGAACACACTTCCTGCATCGCCAAAGAGAT	2325		
QY	1151	GGTCTTCTCGGAATGGCATGAGGACCCACCACTGCTGTGGCCCTGCTGGCCCGCCCTGGG	1220		
Db	2326	CGAGTTTCTGTCGTGGAGCCGAGCCCTTGGGGATCCACACTACAGCTCTTGTGCTGTGGG	2385		
QY	1221	CTTCTCAGACCCCTGGGCGATCTGCTGTATATATCTGGAGGACCTTCAGACCCCATAGT	1280		
Db	2386	CATTTTCTCTACAGCCTTTCGTGCTGGGCGCTCTTTCACAAATTCGCGACACGCCCATCT	2445		
QY	1281	TCGCTCGGCTGGGGGCCCCCATATGCTTCTCGATGCTGACACTGCTGCTGGTGGCATACAT	1340		
Db	2446	CAGGCCACCAACCGGGAGCTCTCTATCTCTTCTTCTTCTCTCTCTCTCTCTCTCTCTCT	2505		
QY	1341	GGTGGTCCCGGTGACGTGGGGGCCGCCCAAGGCTTCACCTGCTGCGCGCCAGGCGCT	1400		
Db	2506	CAGCTCCCTGTTCTTCATCGGGGAGCCCCAGACGTGAGCTGCCGCTCCGCCACCGCGC	2565		
QY	1401	CTTTCCCTCTGCTTACCAATTTGATCTCTCTGTATCTGCCCGTGTCTTTTCCATATGCT	1460		
Db	2566	CTTTGGCATACAGCTCTGCTCTCTGTGATCTCTGATATCTCTGGAAGAAACCAATCGGGTCT	2625		
QY	1461	CTGGGCTTCGAAGATGGGCGAGCGGCTCTCCACAGGCGCTACACACTCTAGGGGTCCGTACCA	1520		
Db	2626	CCTGCTGTTTGA--GGCCAAAGATTCCACACAGCTTCCACCGGAAGTGTGGGGGCTCAA	2682		
QY	1521	GGGCGCTACAGCTCTTATGAGCATTTTATCAGGCTACTCAAAATGGTCATTTGTTGATTTGG	1580		
Db	2683	CCTGCAAGTTCTGCTGGGTCTTCTCTGTGACACTTATGAGAGATTGTACCTGTGGCAATTGG	2742		
QY	1581	CATGCTGACAGCGCTCAATGCTCCACCCCGCTACTGACCCCGATGAGACCCCAAGATCACAAT	1640		

Db 2743 GCTCAATACAGCGCCCTCTGAGCTACCGAACGAGAGGATCATCTT 2802
 QY 1641 TGTCTCTGTAACCCCAATACCGCAACAGCTCTTTCACACAGCCGAGCT 1700
 Db 2803 CATACCTGCGACAGGCGCTGCTCAGCGCGCTTCTCTGATCGGACCTCTT 2862
 QY 1701 GCTCTCAGTGGTGGTTCAGCTTGCCTCATGAGGGAAGAGTGGCCACCACTCAA 1760
 Db 2863 GCTGGCGGCATCTGCTTCTTCTGCTTCAATCCCGAAGTGGCAGAACTTCAA 2922
 QY 1761 CGAGCGCAAGTTCATCACTCCTCAGACATGACCTCTATTTCACTTCCTCTG 1820
 Db 2923 TGAAGCCAGTTCATCACTTCACTCAGATGCTCATCTTCTCTGATCTCTTCA 2982
 QY 1821 CACCTTCATGCTCCACACAGCGGGGCTGAGTCCATGATGAGCTTGGGACCTG 1880
 Db 2983 CCGCGCTACGCGACGACTTACGCAAGTTCGCTGCGCTGAGGATGAGTCCATCT 3042
 QY 1881 GCTCAACCTCTGCGCATACAGCTTGGGCTACTTGGCCCAAGTCAATGATCTCTT 1940
 Db 3043 GCGCGGACGCTTGGCTTGGCTGCGCTGATCTTTCACAGAGTTCATCATCTCTT 3102
 QY 1941 CTACCCGAGCGCAACAGCGCCGCTACTTCAACAGATGATTCAGGCTACACATGAG 2000
 Db 3103 CAAGCTTCCCGAACAACATCGAGAGTGGCTGCGAGCACGCGGACACGCTTCAA 3162
 QY 2001 GAGGAC 2007
 Db 3163 GGTGGC 3169

RESULT 12
 US-08-484-159-1
 ; Sequence 1, Application US/08484159
 ; Patent No. 6313146
 ; GENERAL INFORMATION:
 ; APPLICANT: Bradford C. Van Wagenen
 ; APPLICANT: Manuel F. Balandrin
 ; APPLICANT: Eric G. Del Mar
 ; APPLICANT: Edward F. Nemeth
 ; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: First Interstate World Center
 ; STREET: Suite 4700
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 90071
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: FASTSEQ
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/484,159
 ; FILING DATE: 7 June, 1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; PRIOR APPLICATION DATA: including application
 ; PRIOR APPLICATION DATA: described below: 9
 ; APPLICATION NUMBER: 08/353,784
 ; FILING DATE: 9 December, 1994
 ; APPLICATION NUMBER: PCT/US/94/12117
 ; FILING DATE: 21 October, 1994
 ; APPLICATION NUMBER: U.S. 08/292,827
 ; FILING DATE: 23 August, 1994
 ; APPLICATION NUMBER: U.S. 08/141,248
 ; FILING DATE: 22 October, 1993

; APPLICATION NUMBER: U.S. 08/009,389
 ; FILING DATE: 23 February, 1993
 ; APPLICATION NUMBER: U.S. 08/017,127
 ; FILING DATE: 12 February, 1993
 ; APPLICATION NUMBER: U.S. 07/934,161
 ; FILING DATE: 21 August, 1992
 ; APPLICATION NUMBER: U.S. 07/834,044
 ; FILING DATE: 11 February, 1992
 ; APPLICATION NUMBER: U.S. 07/749,451
 ; FILING DATE: 23 August, 1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Heber, Sheldon O.
 ; REGISTRATION NUMBER: 38,179
 ; REFERENCE/DOCKET NUMBER: 214/101
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (213) 489-1600
 ; TELEFAX: (213) 955-0440
 ; TELEX: 67-3510
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5275 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; MOLECULE TYPE: linear
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 515..3769
 ; OTHER INFORMATION:
 ; US-08-484-159-1

Query Match 8.6%; Score 172.6; DB 4; Length 5275;
 Best Local Similarity 50.2%; Pred. No. 3.4e-28;
 Matches 455; Conservative 0; Mismatches 449; Indels 3; Gaps 1;

QY 1101 TGAATGCCGGAATACAGAGTGGTCTTACAGAGTACCTCTCTTCAAGCGGACGT 1160
 Db 2266 TAAATGCCGGAATACAGAGTGGTCTTACAGAGTACCTCTCTTCAAGCGGACGT 2325
 QY 1161 GGTCTTCCTGGAATGATGATGAGGACGACGACGACGCTGTGGCCCTGTGGCGGCTGGG 1220
 Db 2326 GGAATTTCTGCTGAGACCGAGCCCTTGGGATGCACTACGCTCTTGTGCTGGG 2385
 QY 1221 CTTCCTCAGACCCCTGGGCACTCTGTGATATTCGTGAGGCACTTCCAGACCATATAT 1280
 Db 2386 CATTTCTCAGACCCCTGGGCACTCTGTGATATTCGTGAGGCACTTCCAGACCATATAT 2445
 QY 1281 TCGCTGGGCTGGGCGGCGGCACTGCTTCTGATGCTGACACTGCTGCTGGGATATCAT 1340
 Db 2446 CAAGCGCACCAACCGGAGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2505
 QY 1341 GGTGTCGCGGATGATGAGGCGGCGGCAAGGCTCCACCTGCTGCGGCGGCGCT 1400
 Db 2506 CAGTCTCCCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2565
 QY 1401 CTTCCTCAGACCCCTGGGCACTCTGTGATATTCGTGAGGCACTTCCAGACCATATAT 1460
 Db 2566 CTTGGGATGAGCTTCTGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2625
 QY 1461 CTGGGCTTCAAGATGAGGCGGCGGCTTCCAGCGGCTTCAAGATGAGGCTGCTTCA 1520
 Db 2626 CTGGGCTTCAAGATGAGGCGGCGGCTTCCAGCGGCTTCAAGATGAGGCTGCTTCA 2682
 QY 1521 GGGGCGCTAGCTCTTATGAGGCTTATCAAGGATGAGGCTTATGAGGCTTATGAGGCT 1580
 Db 2683 CTGCACTTCTGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2742
 QY 1581 CATGCTGAGGCGGCTTATGAGGCGGCGGCTTATGAGGCTTATGAGGCTTATGAGGCT 1640
 Db 2743 GCTCAATACAGCGCCCTCTGAGCTACCGAACAGAGCTGAGAGGATCATCTT 2802
 QY 1641 TGTCTCTGTAACCCCAATACCGCAACAGCTGCTTTCACACAGCTTGAAGCTGCT 1700

Db 2803 CATCAGCTGCCAGAGGGCTGCTCATGCGCGCTTCCGTGATCGGCTACACCTGCT 2862
Qy 1701 GCTCAGTGTGGTTTACCTTCGCTTACATGGGCAAAAGCTGCCACCACTAA 1760
Db 2863 GCTGGCCGCAATGCTTCTTCCGCTTCAAGTCCGGAAGCTGCGAGAACTTCAA 2922
Qy 1761 CGAGGCGCAATGCTTACCTTACCTTACCTTATTTTACCTTACCTTCCCTCTG 1820
Db 2923 TGAAGCCAGTTCATCAGCTTACAGATGCTTCTTCTTCAATGCTGATCTTTTAT 2982
Qy 1821 CACCTTACATGCTGCTTACAGCGGGGCTGCTGACCATCTGACCTTGTGACATG 1880
Db 2983 CCGGCTTACGCTTACAGCAATGCTGCTCTGCTGAGGATGATCGCATCT 3042
Qy 1881 GCTCACTGCTGCTTACAGCTTGGCTTACCTTGGCGCCCAAGTCTTACATGCTCT 1940
Db 3043 GCGGCGCACTTGGCTTGGCTTGGCTTGTATTTCTTCAACAGGCTTACATCATCTCTT 3102
Qy 1941 CTACCCGAGCGCAACAGCCCGCTTACTTACAGCATGATCCAGGCTTACCATGAG 2000
Db 3103 CAAGCTTCCGGAACACATCATGAGAGGTGCGTGCAGCAACCGGGGCAACGCTTCAA 3162
Qy 2001 GAGGAC 2007
Db 3163 GGTGGC 3169

RESULT 13

Sequence 3, Application US/08485588

Patent No. 5688938

GENERAL INFORMATION:

APPLICANT: Edward M. Brown

APPLICANT: Steven C. Hebert

APPLICANT: Forrest H. Fuller

TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE

TITLE OF INVENTION: MOLECULES

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: First Interstate World Center

STREET: Suite 4700

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,588

FILING DATE: 7 June, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA: including application

PRIOR APPLICATION DATA: described below: 9

APPLICATION NUMBER: 08/353,784

FILING DATE: 9 December, 1994

APPLICATION NUMBER: PCT/US/94/12117

FILING DATE: 21 October, 1994

APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994

APPLICATION NUMBER: U.S. 08/141,248

FILING DATE: 22 October, 1993

APPLICATION NUMBER: U.S. 08/009,389

FILING DATE: 23 February, 1993

APPLICATION NUMBER: U.S. 08/017,127

FILING DATE: 12 February, 1993

APPLICATION NUMBER: U.S. 07/934,161

FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
SEQUENCE CHARACTERISTICS:
LENGTH: 3809 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 373..3606
OTHER INFORMATION:
US-08-485-588-3

Query Match 8.3%; Score 167; DB 1; Length 3809;

Best Local Similarity 49.8%; Pred. No. 5e-27;

Matches 450; Conservative 0; Mismatches 450; Indels 3; Gaps 1;

Qy 1103 AATGCCGAATACAGAGTGTCTTACAGAGTGAAGCTCTCTTCAAGCGGACGTGG 1162
Db 2123 AGTGCCAGATGACTTCTGTGTCATGAGAAACACACCTCTGTGATTCAGAGGATCG 2182
Qy 1163 TCTTCTGGAATGGCATGAGGACCCACCATGCTGTGGCCCGCCCTGGGCT 1222
Db 2183 AGTTCTGTGTCGACGAGGACCTTGGGATGCGACTCATCACCCTTTCGCGTGGGCA 2242
Qy 1223 TCTTCAGACCTTGGGACCTCTGTGATATCTGGAGGACTTCCAGACCCATAGTTC 1282
Db 2243 TTTTCTGACAGACCTTGTGTGGGTGTATATCAAGTTCCGCAACACCATATGCA 2302
Qy 1283 GCTGGCTGGGGGCCCATGTGCTCTGATGCTGACACTGCTGTGGCATGATG 1342
Db 2303 AGGCCACCAACCGAGAGCTCTCTTACCTCTCTCTCTCTCTCTCTCTCTCTCA 2362
Qy 1343 TGTGCCGCTGTACGTGGGGGCCCAAGTCTCCACCTGCTGCGGACGCGCTCT 1402
Db 2363 GCTCCCTGTCTTCAATCGGGAGCCCGAGACTGAGCTGCGCGCTGCGGCGGCT 2422
Qy 1403 TTCCCTGCTTACATTTGCAATTCATCTGATGCGGCTGCTTCTTCCAGATGCTCT 1462
Db 2423 TTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2482
Qy 1463 GCGCTTCAAGATGCGCCAGCGCTTCCAGCGGCTTACAGTACTGCTGCTGCTGCT 1522
Db 2483 TGTGTTTGA---GGCCAAAGTCCCGACAGCTTCCAGCGAGTGGGGGCTCAAC 2539
Qy 1523 GCGCTTACGCTCTATGCGCTTATACAGGCTACTCAAAATGCTATTTGTAATGCA 1582
Db 2540 TGCAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2599
Qy 1583 TGTGCGCAGCGCTTACAGTCCAGCCCGCTTACGAGCCGAGTACCCCAAGTACCAAT 1642
Db 2600 TCTACACCGCGCCCGCTTCAAGTACGCAACGAGCTGAGAGTACGATCACTTCA 2659
Qy 1643 TCTCTTACACCGCAATACGCAACGAGCTGCTGCTTCAACAGCAGCTGAGCTGCTG 1702
Db 2660 TCAAGTGCAGAGAGGCTCCCTCAATGCGGCTTCTGATGCGTACACGCTGCTG 2719
Qy 1703 TCTAGTGTGCTTACGCTTCCGCTTACATGAGGCAAAAGCTGCGCCACCACTGCAAG 1762
Db 2720 TGGCTGCATCTGCTTCTTCTTGTGCTTCAAGTCCCGGAGCTGCGGAGAACTTCAATG 2779

Db 2840 CAGCCTATGACGACACCTATGAGTGTGTCTGTGCGGTAGAGTGATGTCATCCGCG 2899
Qy 1883 TCACCTCTCTGACATGACCTTGCGCTACTTGCGCCCAAGTGATGATCTCTCT 1942
Db 2900 CAGCAGCTTGTGGCTGTGGCGGTCTCTTCAACAGATCTACATCTCTCTCA 2959
Qy 1943 ACCCGAGGAGGAGGCGCGCTTCAACAGCATGATCCAGGCGTACACCATGAGA 2002
Db 2960 AGCCATCCCGACACCATGAGAGGTGCGTTGACGACCGACCTCAGCTTCAAGG 3019
Qy 2003 GGG 2005
Db 3020 TGG 3022

RESULT 15
US-08-480-751-3
Sequence 3, Application US/08480751
Patent No. 5858684

GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,751
FILING DATE: 7 June, 1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/84/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991

ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/004
TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3809 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 373..3606
OTHER INFORMATION:
US-08-480-751-3

Query Match 8.3%, Score 167; DB 2; Length 3809;
Best Local Similarity 49.8%; Pred. No. 5e-27;
Matches 450; Conservative 0; Mismatches 450; Indels 3; Gaps 1;

Qy 1103 AATGCCGATATACGAGTGTCTTACAGAGAGACCTCTGCTTCAAGGCGGACGTGG 1162
Db 2123 AGTGCCAGATGACTTGTGCTCAATGAGAACACCTCTGCTTCAAGGAGATCG 2182
Qy 1163 TCTTCGTAATGGCATGAGGACACCATGCTGTGGCCCTGCTGGCCGCGGCGT 1222
Db 2183 AGTTCTGTGCGAGGAGGCGCTTGGAGATGCACTACCTCTTGTGCGCTGCGCA 2242
Qy 1223 TCTTCGAGACCCCTGACATCTGTGATATCTGAGGCACTTCCAGACCATAGTTC 1282
Db 2243 TTTTCTGACACCTTGTGCGGTGTGTTATCAAGTTCGCAACACCATATGCA 2302
Qy 1283 GCTGCGTGGGGGCCCATGTGCTTCTGATGCTACACTGTGCTGTGGATCATGG 1342
Db 2303 AGGCCACCAACCGAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 2362
Qy 1343 TGTGCGGATGACGTGGGGGCGCCCAAGGCTCCACCTGCGCGCGAGGCGCT 1402
Db 2363 GCTGCTGTCTTCAATCGGGGAGGCGCCAGACTGAGACCTGCGCGCGAGGCGCT 2422
Qy 1403 TTCCCTCTGCTTCAAAATTCATCTCTGTATGCGCGGCTTCTTCCAGATGCT 1462
Db 2423 TTGGCATGAGCTTCTGCTGCTCATGATGATGATGATGATGATGATGATGATGAT 2482
Qy 1463 GCGCCTTCAAGATGCGGCGGCTTCCAGCGGCTTCAAGTACGACTGAGGCGCTCAAGG 1522
Db 2483 TGTGTTTGA---GGCCAGATCCCGACAGCTTCCACCGCAAGTGTGGGGGCTCAAC 2539
Qy 1523 GGCCTACGCTCTATGAGCTTATACAGGATCACTCAAAATGCTATGTTGTAATGGCA 1582
Db 2540 TGCAATTCCTGTGTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2599
Qy 1583 TGCTGGACGCGCTTACGCTTCCAGCGGCTTACGCTGAGCCGATGACCCAGATCAAAATG 1642
Db 2600 TCTACACGCGCGCGCGCTTCAAGCTTACGAGAGATGAGATGATGATGATGATGAT 2659
Qy 1643 TCTCTGTAACCCCAATACCGCAACAGCTCTGCTTCAAGCAAGCTGAGCTGCTGCT 1702
Db 2660 TCACGTGCAAGAGGCGCTTCCATGAGCGCTGAGGCTTCTGATGCGCTACACCTGCTG 2719
Qy 1703 TCTCAGTGTGTGCTTCAAGCTTGTGCTTCAAGGAGAGGCTGCGCAACCTACAGAG 1762
Db 2720 TGGCGCATGCTGCTTCTTGTGCTTCAAGTCCGGAAGTGGCGGAGATCTTCAAG 2779
Qy 1763 AGCCCAAGTATACACCTGAGATGACCTTCTATTTCACTCACTGCTGCTGCTGCA 1822
Db 2780 AAGCCAAAGTATACACCTGAGATGATGATGATGATGATGATGATGATGATGATG 2839
Qy 1823 CTTGATGCTGCTACAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1882
Db 2840 CAGCCTATGACGACACCTATGAGAGTGTGTGTGCGGTAGAGTGATGTCATCCGCG 2899
Qy 1883 TCACCTCTCTGACATGACCTTGCGCTACTTGCGCCCAAGTGATGATGATCTCTCT 1942

```
Db 2900 CAGCCAGCTTTGGCTTGCTGCGCTGCACTCTTCAACAGATCTACATCATCTCTTCA 2959
QY 1943 ACCCGAGCGCAACAGCGCCCTACTTCAACAGCATGATCCAGGCTACACCATGAGA 2002
Db 2960 AGCCATCCCGCAACACCATGAGGAGGTGCGTTGCAAGCACCGCAGCTCACGCTTTCAAG 3019
QY 2003 GGG 2005
Db 3020 TGG 3022
```

Search completed: May 23, 2003, 18:23:42
Job time : 156.947 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 12:46:24 ; Search time 248.116 Seconds
(Without alignments)
10697.149 Million cell updates/sec

Title: US-09-927-315-12

Perfect score: 2010

Sequence: 1 ataccctacagcgcctacag.....acaccatgagggagactag 2010

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PC1_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PC1US_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2010	100.0	2010	9	US-10-096-144-8
2	2010	100.0	2010	9	US-09-927-315-12
3	2010	100.0	2010	12	US-10-124-598-8
4	1928.6	96.0	2520	9	US-09-897-427A-3
5	1928.6	96.0	2520	9	US-10-035-045-23
6	1232.8	61.3	2993	9	US-10-036-144-3
7	1232.8	61.3	2993	9	US-09-927-315-10
8	1232.8	60.3	2993	12	US-10-124-598-3
9	1212.6	60.3	2532	9	US-10-096-144-4
10	1212.6	60.3	2532	9	US-09-927-315-11
11	1212.6	60.3	2532	12	US-10-124-598-4
12	392.8	19.5	2771	9	US-08-361-652-4
13	352.8	19.5	2771	9	US-09-927-315-4
14	386.4	19.2	2292	10	US-09-819-946-3
15	386.4	19.2	2526	10	US-09-819-946-3
16	386.4	19.2	2951	10	US-09-819-946-9
17	385.2	19.2	2526	9	US-09-897-427A-1
18	385.2	19.2	2526	9	US-09-799-629-16
19	385.2	19.2	2526	9	US-10-035-045-16

20	376.6	18.7	2333	9	US-09-361-652-6	Sequence 6, Appl1
21	376.6	18.7	2333	9	US-09-927-315-6	Sequence 6, Appl1
22	359.2	17.9	2579	9	US-09-361-652-5	Sequence 5, Appl1
23	359.2	17.9	2579	9	US-09-927-315-5	Sequence 5, Appl1
24	258	12.8	697	10	US-09-811-284-117	Sequence 117, App
25	218.6	10.9	2559	9	US-09-897-427A-5	Sequence 5, Appl1
26	218.6	10.9	2559	9	US-09-897-427A-7	Sequence 7, Appl1
27	218.6	10.9	2559	9	US-09-927-315-14	Sequence 14, Appl1
28	218.6	10.9	2559	9	US-10-035-045-3	Sequence 3, Appl1
29	218.6	10.9	3458	9	US-10-261-482-1	Sequence 1, Appl1
30	217	10.8	2559	10	US-10-282-837-15	Sequence 15, Appl1
31	217	10.8	2559	10	US-09-786-338A-15	Sequence 15, Appl1
32	217	10.8	3489	10	US-10-282-837-13	Sequence 13, Appl1
33	217	10.8	3489	10	US-09-786-338A-13	Sequence 13, Appl1
34	216.2	10.8	2553	9	US-09-799-629-3	Sequence 3, Appl1
35	190.4	9.5	8194	9	US-09-799-629-15	Sequence 15, Appl1
36	190.4	9.5	8194	9	US-10-035-045-15	Sequence 15, Appl1
37	187	9.3	4134	9	US-09-975-553-17	Sequence 17, Appl1
38	187	9.3	4134	9	US-10-125-792-1	Sequence 1, Appl1
39	187	9.3	4134	9	US-10-125-778-1	Sequence 1, Appl1
40	187	9.3	4134	9	US-10-270-785-17	Sequence 17, Appl1
41	184.2	9.2	1101	10	US-09-819-946-5	Sequence 5, Appl1
42	167	8.3	3361	12	US-10-002-854-1	Sequence 1, Appl1
43	159	7.9	3237	10	US-09-727-205-1	Sequence 7, Appl1
44	157	7.8	705	10	US-09-819-946-7	Sequence 7, Appl1
45	127.4	6.3	2577	9	US-09-799-629-13	Sequence 13, Appl1

ALIGNMENTS

RESULT 1
US-10-096-144-8
Sequence 8, Application US/10096144
Publication No. US20030022288A1
GENERAL INFORMATION:
APPLICANT: Zuker, Charles S.
APPLICANT: Adler, Jon Elliott
APPLICANT: Lindemeyer, Juergen
TITLE OF INVENTION: The Regents of the University of California
TITLE OF INVENTION: Involved in Sensory Transduction
FILE REFERENCE: 023072-088720US
CURRENT APPLICATION NUMBER: US/10/096,144
CURRENT FILING DATE: 2002-03-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/361,631
PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/095,464
PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 2010
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human G-protein coupled receptor (GPCR) B4
US-10-096-144-8
Query Match 100.0%; Score 2010; DB 9; Length 2010;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCACTACAGCGCCATCAGCGATGAGTGGAGACAAGAGTCCCGCTTGGCTG 60
DB 1 ATCACTACAGCGCCATCAGCGATGAGTGGAGACAAGAGTCCCGCTTGGCTG 60
QY 61 CGTACCAACCCAGCGCCGACCACTGCGAGCGGCGGCGGCGGCGGCGGCGGCGG 120
DB 61 CGTACCAACCCAGCGCCGACCACTGCGAGCGGCGGCGGCGGCGGCGGCGGCGG 120

```

QY 121 CGGTGAACTGATCATTTGCTGTGTGAGCAGCAGACCTATGGCCGCGACAATGGCCAG 180
    |||||
Db 121 CGGTGAACTGATCATTTGCTGTGTGAGCAGCAGACCTATGGCCGCGACAATGGCCAG 180
QY 181 CTGCTTGGCAGAGCGCGTGGCCCGGCGACATCTGCATCGCTCTTCCAGAGAGCGTGGC 240
    |||||
Db 181 CTGCTTGGCAGAGCGCGTGGCCCGGCGACATCTGCATCGCTCTTCCAGAGAGCGTGGC 240
QY 241 ACAGTGCAGCCCAACGACATGAGTGCAGAGAGAGCGCGCGCGGCGATGATG 300
    |||||
Db 241 ACAGTGCAGCCCAACGACATGAGTGCAGAGAGAGCGCGCGCGGCGATGATG 300
QY 301 GACAGCTGAGCAGAGACAGCAGCGCGCTGTGCTGTCTTCGCGCCGACCTGACCTG 360
    |||||
Db 301 GACAGCTGAGCAGAGACAGCAGCGCGCTGTGCTGTCTTCGCGCCGACCTGACCTG 360
QY 361 TACCACTTCTTCAATGAGTGTGTGCGCAGAACTTCAAGGCGCGCGTGTGATCGCTCC 420
    |||||
Db 361 TACCACTTCTTCAATGAGTGTGTGCGCAGAACTTCAAGGCGCGCGTGTGATCGCTCC 420
QY 421 GAGTCTGGGCGATGACAGCGCGGTGCTGACAACTCAGAGAGCGTGGCGCACTTGGGAGCC 480
    |||||
Db 421 GAGTCTGGGCGATGACAGCGCGGTGCTGACAACTCAGAGAGCGTGGCGCACTTGGGAGCC 480
QY 481 TTCTGGGCGATCAACATCAAGAGCGTCCATCCGCGCTTCAAGTGAATTCGCGAGTGG 540
    |||||
Db 481 TTCTGGGCGATCAACATCAAGAGCGTCCATCCGCGCTTCAAGTGAATTCGCGAGTGG 540
QY 541 GGGCCACAGGCTGGGCGCGCCACCTTCAGAGAGAGCGAGAGCTTACCTGACAGCAG 600
    |||||
Db 541 GGGCCACAGGCTGGGCGCGCCACCTTCAGAGAGAGCGAGAGCTTACCTGACAGCAG 600
QY 601 GAGTGCAGCAACTGCTGGAAGCGCACCTTGCTTCAACACCATTTCAAGGCTCTGGG 660
    |||||
Db 601 GAGTGCAGCAACTGCTGGAAGCGCACCTTGCTTCAACACCATTTCAAGGCTCTGGG 660
QY 661 GAGGTGTGCTTACAGAGCTGTACTGTGCGGTCTATGCTGTGGCCCATGCGCTGACAGC 720
    |||||
Db 661 GAGGTGTGCTTACAGAGCTGTACTGTGCGGTCTATGCTGTGGCCCATGCGCTGACAGC 720
QY 721 CTCTCTGGGCTGTGACAAAGACAGCTGACCAAGAGGCTGTCTACCCCTGGAGCTGCTT 780
    |||||
Db 721 CTCTCTGGGCTGTGACAAAGACAGCTGACCAAGAGGCTGTCTACCCCTGGAGCTGCTT 780
QY 781 GAGGAGATCTGGAAGGCTCACTTCACTCTCTGAGACCAACAAATCTTCTTGACCCGCAA 840
    |||||
Db 781 GAGGAGATCTGGAAGGCTCACTTCACTCTCTGAGACCAACAAATCTTCTTGACCCGCAA 840
QY 841 GGGGAGCTGGCTGTGACCTTGGAGATTGTCAAGTGGACCGGAGCCAGAAATCCC 900
    |||||
Db 841 GGGGAGCTGGCTGTGACCTTGGAGATTGTCAAGTGGACCGGAGCCAGAAATCCC 900
QY 901 TTCCAGAGCGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
    |||||
Db 901 TTCCAGAGCGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
QY 961 CTGACACCGCTTCAACACAGATGCTGATGCTATGCTTCAAGAGGCTGACGAGTGGG 1020
    |||||
Db 961 CTGACACCGCTTCAACACAGATGCTGATGCTATGCTTCAAGAGGCTGACGAGTGGG 1020
QY 1021 CAAAAGAGAGGCTGTGGGCTATCAGCTGTGCTTCAAGTGCATGAGTGGCTTCCC 1080
    |||||
Db 1021 CAAAAGAGAGGCTGTGGGCTATCAGCTGTGCTTCAAGTGCATGAGTGGCTTCCC 1080
QY 1081 GGCACCTTCTCTCAACACAGCTGAATGCCGAATPACAGAGTGTCTTCAAGAGTGGAGCC 1140
    |||||
Db 1081 GGCACCTTCTCTCAACACAGCTGAATGCCGAATPACAGAGTGTCTTCAAGAGTGGAGCC 1140
QY 1141 TTCTGTCTTCAAGGCGAGCTGTGCTTCTCTGAGATGAGGAGGAGGAGGAGGAGGAGTGG 1200
    |||||
Db 1141 TTCTGTCTTCAAGGCGAGCTGTGCTTCTCTGAGATGAGGAGGAGGAGGAGGAGGAGTGG 1200

```

```

QY 1201 GCCCTGTGGCCGCGCTTGGGCTTCTTCAGACACCTTGCGCATCTCTGATATTTCTGGAG 1260
    |||||
Db 1201 GCCCTGTGGCCGCGCTTGGGCTTCTTCAGACACCTTGCGCATCTCTGATATTTCTGGAG 1260
QY 1261 CACTTCCAGACACCCATAGTTCGCTCGGCTGGGGGCCCATGTGTCTTCTGATGTGACA 1320
    |||||
Db 1261 CACTTCCAGACACCCATAGTTCGCTCGGCTGGGGGCCCATGTGTCTTCTGATGTGACA 1320
QY 1321 CTGCTGTGGGCGATGATGATGTGTGTCGCGGTGATGATGAGTGGGGGCCCAAGGCTTCCAC 1380
    |||||
Db 1321 CTGCTGTGGGCGATGATGATGTGTGTCGCGGTGATGATGAGTGGGGGCCCAAGGCTTCCAC 1380
QY 1381 TGCCTTGCAGCCAGAGGCGCTTCTTCCCTCTGCTTCAAAATTTGACATCTCTGATGCC 1440
    |||||
Db 1381 TGCCTTGCAGCCAGAGGCGCTTCTTCCCTCTGCTTCAAAATTTGACATCTCTGATGCC 1440
QY 1441 GTGCGTCTTCCAGATGCTGTGCGCTTCAAGATGAGGCGCGCTTCCAGCGGCTTAC 1500
    |||||
Db 1441 GTGCGTCTTCCAGATGCTGTGCGCTTCAAGATGAGGCGCGCTTCCAGCGGCTTAC 1500
QY 1501 AGCTACTGGGTCGCGCTACAGAGGCGCGCTTACGCTCTATGAGCATTTATCAGGTAATCAA 1560
    |||||
Db 1501 AGCTACTGGGTCGCGCTACAGAGGCGCGCTTACGCTCTATGAGCATTTATCAGGTAATCAA 1560
QY 1561 ATGCTATGTGTAAATTTGAGATGTGGCAGCGGCGCTTACGCGCTTACAGCCCGTACTGACCC 1620
    |||||
Db 1561 ATGCTATGTGTAAATTTGAGATGTGGCAGCGGCGCTTACGCGCTTACAGCCCGTACTGACCC 1620
QY 1621 GATGACCCCAAGATCACAAATTTCTCTGTAAACCCCACTACGCAACACGCTGTGTTTC 1680
    |||||
Db 1621 GATGACCCCAAGATCACAAATTTCTCTGTAAACCCCACTACGCAACACGCTGTGTTTC 1680
QY 1681 AACACAGCGCTTGGACCTGCTGCTCTCAAGTGGGTTTCAAGCTTGGCTATGAGGCAAA 1740
    |||||
Db 1681 AACACAGCGCTTGGACCTGCTGCTCTCAAGTGGGTTTCAAGCTTGGCTATGAGGCAAA 1740
QY 1741 GAGCTGCGCACCAACTTCAACAGAGGCGCAAGTTATCAACCTCAGACGATGACCTTATTTTC 1800
    |||||
Db 1741 GAGCTGCGCACCAACTTCAACAGAGGCGCAAGTTATCAACCTCAGACGATGACCTTATTTTC 1800
QY 1801 ACCTCATCGCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860
    |||||
Db 1801 ACCTCATCGCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860
QY 1861 GTGGAGCTTGTGGATGCTGCTCAACCTCTGCGCATCAGGCTGAGGCTTCTGCGGCC 1920
    |||||
Db 1861 GTGGAGCTTGTGGATGCTGCTCAACCTCTGCGCATCAGGCTGAGGCTTCTGCGGCC 1920
QY 1921 AAGTCTACATGATCTCTTCTTACCGGAGCGCAACAGCGCGCTTACTTCAACAGCATG 1980
    |||||
Db 1921 AAGTCTACATGATCTCTTCTTACCGGAGCGCAACAGCGCGCTTACTTCAACAGCATG 1980
QY 1981 ATCCAGGCGTACACCATGAGAGGAGACTAG 2010
    |||||
Db 1981 ATCCAGGCGTACACCATGAGAGGAGACTAG 2010

```

RESULT 2
 US-09-927-315-12
 ; Sequence 12, Application US/09927315
 ; Publication No. US2003004045A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zuker, Charles S.
 ; APPLICANT: Ryba, Nicholas J.P.
 ; APPLICANT: Nelson, Greg
 ; APPLICANT: Hoon, Mark A.
 ; APPLICANT: Chandrasekhar, Jayaram
 ; APPLICANT: Zhang, Yifeng
 ; APPLICANT: The Regents of the University of California
 ; APPLICANT: as represented by the Secretary of the
 ; Department of Health and Human Services
 ; TITLE OF INVENTION: Mammalian Sweet Taste Receptors

FILE REFERENCE: 02307E-120110US
CURRENT APPLICATION NUMBER: US/09/927,315
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: US 60/302,898
PRIOR FILING DATE: 2001-07-03
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 2010
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human T1R2 sweet taste receptor
US-09-927-315-12

Query Match 100.0%; Score 2010; DB 9; Length 2010;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCACTACAGCGGCATGAGGATGAGTGGAGACAAGGTCGCTCCGGCTTGGTG 60
DB 1 ATCACTACAGCGGCATGAGGATGAGTGGAGACAAGGTCGCTCCGGCTTGGTG 60
QY 61 CGTACACACCCAGCGCGACACACGCTGAGGGCATGTGTGACACTGATGCTGCACTTC 120
DB 61 CGTACACACCCAGCGCGACACACGCTGAGGGCATGTGTGACACTGATGCTGCACTTC 120
QY 121 CGCTGGAACTGATATATTTGTGTGTGTGAGACGACGACCTATATGCGCGACAAATGGCCAG 180
DB 121 CGCTGGAACTGATATATTTGTGTGTGTGAGACGACGACCTATATGCGCGACAAATGGCCAG 180
QY 181 CTGCTTGGCGAGCGCGGTGGCGCGGCGAGACATCTGATCGCTTCCAGAGAGCGGTGGCC 240
DB 181 CTGCTTGGCGAGCGCGGTGGCGCGGCGAGACATCTGATCTGCTTCCAGAGAGCGGTGGCC 240
QY 241 ACACGTGACCCCAACACAGATGACGTGAGAGGCGCCAGCGCTGTGTGACATTTGTG 300
DB 241 ACACGTGACCCCAACACAGATGACGTGAGAGGCGCCAGCGCTGTGTGACATTTGTG 300
QY 301 GACAAAGTCGACGAGAGACGAGCGCGGTGTCGTGTGTTCGCCCGACCTGACCCCTG 360
DB 301 GACAAAGTCGACGAGAGACGAGCGCGGTGTCGTGTGTTCGCCCGACCTGACCCCTG 360
QY 361 TACCACTTCTCAATGAGGTGTGTGCGCAGAACTTCAGCGGGCGCGTGTGATTCCTCC 420
DB 361 TACCACTTCTCAATGAGGTGTGTGCGCAGAACTTCAGCGGGCGCGTGTGATTCCTCC 420
QY 421 GAGTCTGGGCCATGACCCCGGTCTGTGCAAACTTCACGAGAGTGGCCACTTGGGCACC 480
DB 421 GAGTCTGGGCCATGACCCCGGTCTGTGCAAACTTCACGAGAGTGGCCACTTGGGCACC 480
QY 481 TTCTGGGCGCATCAATCAAGAGCTGGCCATCCCGGGCTTCAGTGAATTCGCGCAATGG 540
DB 481 TTCTGGGCGCATCAATCAAGAGCTGGCCATCCCGGGCTTCAGTGAATTCGCGCAATGG 540
QY 541 GGCCACAGGCTGGGCGCGCCACCCCTCAGAGAGACGACGACGATATACCTGCAACGAG 600
DB 541 GGCCACAGGCTGGGCGCGCCACCCCTCAGAGAGACGACGACGATATACCTGCAACGAG 600
QY 601 GAGTGCAGCAACTGCTGAGAGCGCACTGTTCCTTCAACACCAATTCAGAGGCTCTG 660
DB 601 GAGTGCAGCAACTGCTGAGAGCGCACTGTTCCTTCAACACCAATTCAGAGGCTCTG 660
QY 661 GAGCGTGTCTTACAGGCTGTACTGTGGGTCTATGCTGTGGCCCATGCGCTGACACAG 720
DB 661 GAGCGTGTCTTACAGGCTGTACTGTGGGTCTATGCTGTGGCCCATGCGCTGACACAG 720
QY 721 CTCCTGCGGTGTGACAAAAGCACTGACCAAGAGGGGTGTACCCCTGGAGAGCTGCT 780
DB 721 CTCCTGCGGTGTGACAAAAGCACTGACCAAGAGGGGTGTACCCCTGGAGAGCTGCT 780
QY 781 GAGGAGATCTGGAAGTCAACTTCACTCTCTGAGACCAATATCTTTCGACCCGCA 840
DB 781 GAGGAGATCTGGAAGTCAACTTCACTCTCTGAGACCAATATCTTTCGACCCGCA 840

DB 781 GAGGAGATCTGGAAGTCAACTTCACTCTCTGAGACCAATATCTTTCGACCCGCA 840
QY 841 GGGGAGGTGGCTCTGCACTTGGAGATTGTCCAGTGGCAATGGAGCCGAGATCC 900
DB 841 GGGGAGGTGGCTCTGCACTTGGAGATTGTCCAGTGGCAATGGAGCCGAGATCC 900
QY 901 TTCCAGAGGTGGCTCTGCACTTGGAGATTGTCCAGTGGCAATGGAGCCGAGATCC 960
DB 901 TTCCAGAGGTGGCTCTGCACTTGGAGATTGTCCAGTGGCAATGGAGCCGAGATCC 960
QY 961 CTGCACACCGTTCACACACGATCCCATATGTCATGTGTTCACAGAGTGGCCAGTCAAGG 1020
DB 961 CTGCACACCGTTCACACACGATCCCATATGTCATGTGTTCACAGAGTGGCCAGTCAAGG 1020
QY 1021 CAAAAGAAAGACCTGTGGGATCCAGCTGTCTGTCTGAGTGCATGATGCTCTCC 1080
DB 1021 CAAAAGAAAGACCTGTGGGATCCAGCTGTCTGTCTGAGTGCATGATGCTCTCC 1080
QY 1081 GGCACCTTCCTCAACACACGATGATGCGAATTAACGATGATGCTCTACAGATGAGACC 1140
DB 1081 GGCACCTTCCTCAACACACGATGATGCGAATTAACGATGATGCTCTACAGATGAGACC 1140
QY 1141 TCTGCTTCAGCGGCGAGCTGTCTTCTGTGAATGAGCATGAGGACCCACATCGCTGTG 1200
DB 1141 TCTGCTTCAGCGGCGAGCTGTCTTCTGTGAATGAGCATGAGGACCCACATCGCTGTG 1200
QY 1201 GCGCTGTGGCGCGCTGGGCTTCTCAGACACCTGTGGCATCTGTGTATTTCTGGAGG 1260
DB 1201 GCGCTGTGGCGCGCTGGGCTTCTCAGACACCTGTGGCATCTGTGTATTTCTGGAGG 1260
QY 1261 CACTTCCAGACACCCATGATGTCGCGGTGGGGGCCCATGTGCTCTGTATGTGTGACA 1320
DB 1261 CACTTCCAGACACCCATGATGTCGCGGTGGGGGCCCATGTGCTCTGTATGTGTGACA 1320
QY 1321 CTGCTGTGTGTGATATATGATGTCGCGGTGGGGGCCCATGTGCTCTGTATGTGTGACA 1380
DB 1321 CTGCTGTGTGTGATATATGATGTCGCGGTGGGGGCCCATGTGCTCTGTATGTGTGACA 1380
QY 1381 TGGCTCTGCGCGCAGGCGCTCTTCCCTGTGCTTCAAAATTTGATGTCCTGTATGCGC 1440
DB 1381 TGGCTCTGCGCGCAGGCGCTCTTCCCTGTGCTTCAAAATTTGATGTCCTGTATGCGC 1440
QY 1441 GTGCGTTCCTTCCAGATGTCGTGCGCTTCAAGATGGCGCGGCTTCCACGCGCCTAC 1500
DB 1441 GTGCGTTCCTTCCAGATGTCGTGCGCTTCAAGATGGCGCGGCTTCCACGCGCCTAC 1500
QY 1501 AGCTACTGGGTCCGCTACAGGAGGCGCTTACGTCCTATGCGATTTATCAGGTACTCAA 1560
DB 1501 AGCTACTGGGTCCGCTACAGGAGGCGCTTACGTCCTATGCGATTTATCAGGTACTCAA 1560
QY 1561 ATGGTCATTTGTTAATGGCATGGTGGGACGCGCTCAGTCCACCGCGTACTGACCCC 1620
DB 1561 ATGGTCATTTGTTAATGGCATGGTGGGACGCGCTCAGTCCACCGCGTACTGACCCC 1620
QY 1621 GATGACCCCAATATCAATTTGTCTGTATACCCCAATACCGCAACAGCCTGTCTTC 1680
DB 1621 GATGACCCCAATATCAATTTGTCTGTATACCCCAATACCGCAACAGCCTGTCTTC 1680
QY 1681 AACACCAAGCTGGACCTGCTCTCATAGTGTGGTTTCAGCTTGGCTTACATGGGCAAA 1740
DB 1681 AACACCAAGCTGGACCTGCTCTCATAGTGTGGTTTCAGCTTGGCTTACATGGGCAAA 1740
QY 1741 GAGCTGCCCACTACACAGAGGCAAGTTCATACACCTCAGCATGACATCTATTTTC 1800
DB 1741 GAGCTGCCCACTACACAGAGGCAAGTTCATACACCTCAGCATGACATCTATTTTC 1800
QY 1801 ACCTATCCGTCCTCCCTGCACTTCATGTGCTTACAGGCGGGGTGTGCTGACATTC 1860
DB 1801 ACCTATCCGTCCTCCCTGCACTTCATGTGCTTACAGGCGGGGTGTGCTGACATTC 1860
QY 1861 GTGAGACTTGTGTGATCACTGTGCTCAACCTCTGAGCATGAGCTGTGCTTGGCCCC 1920
DB 1861 GTGAGACTTGTGTGATCACTGTGCTCAACCTCTGAGCATGAGCTGTGCTTGGCCCC 1920


```

Db 1444 TCCTGGCACACGCTCAACAGACGATCCCTATGCTCATGTGTTCAGAGGTCACAGTCA 1503
Oy 1018 GGGCAAAAGAGAGACCGCTGTGGGATCCACGCTGCTGCTTGCATGTCATGTCCTT 1077
Db 1504 GGGCAAAAGAGAGAACCGTGTGGGATCCACGCTGCTGCTTGCATGTCATGTCCTT 1563
Oy 1078 CCGGACCTTCTTCAACCACTGAA-----TGGCCGAATTAAC 1116
Db 1504 CCGGACCTTCTTCAACCACTGAAATGAATGAATGTCAGGCTGCTGCTGGAATTAAC 1623
Oy 1117 GAGTGTCTTACAGAGTGAAGCTCTGCTTCAAGCGGACGCTGCTTCTTCTGGAATG 1176
Db 1624 GAGTGTCTTACAGAGTGAAGCTCTGCTTCAAGCGGACGCTGCTTCTTCTGGAATG 1683
Oy 1177 CATGAGCACCCACCATGCTGTGGCCCTGTGGCCCTGTGGCTTCTTCTGACACCTG 1236
Db 1684 CATGAGGACCCACCATGCTGTGGCCCTGTGGCCCTGTGGCTTCTTCTGACACCTG 1743
Oy 1237 GCGATCTGTGTATTTCTGGAGGACATTCAGACACCATGCTGCTGCTGCTGGGGG 1296
Db 1744 GCGATCTGTGTATTTCTGGAGGACATTCAGACACCATGCTGCTGCTGCTGGGGG 1803
Oy 1297 CCGATGTCTTCTTCTGATGCTGACACTGCTGTGTGTGGCATATACATGCTGCTGCTAC 1356
Db 1804 CCGATGTCTTCTTCTGATGCTGACACTGCTGTGTGTGGCATATACATGCTGCTGCTAC 1863
Oy 1357 GTGGGGGCGCCCAAGGCTCCACCTGCTTCCGCCAGAGCCCTTTCCTTCTGCTTC 1416
Db 1864 GTGGGGGCGCCCAAGGCTCCACCTGCTTCCGCCAGAGCCCTTTCCTTCTGCTTC 1923
Oy 1417 ACAATTTGATCTCTCTGTATATGCGGCTTCTTCAAGATGCTGTGGCCCTTCAAGATG 1476
Db 1924 ACAATTTGATCTCTCTGTATATGCGGCTTCTTCAAGATGCTGTGGCCCTTCAAGATG 1983
Oy 1477 GCGACCGCTTCCACAGGCGCTACAGTACTGTGGTCCGCTACAGAGGCGCTTCTCT 1536
Db 1984 GCGACCGCTTCCACAGGCGCTACAGTACTGTGGTCCGCTACAGAGGCGCTTCTCT 2043
Oy 1537 ATGGAATTTATACAGGTAATCAAAATGTCATGTTGTAATGTCATGCTGCTGCTG 1594
Db 2044 ATGGAATTTATACAGGTAATCAAAATGTCATGTTGTAATGTCATGCTGCTGCTG 2103
Oy 1595 CTGATCTCCACC--CCGTAATGACCCGATGACCCCAAGATACAAATGTCCTGTAAC 1653
Db 2104 CTGATCTCCACCACCCGTAATGACCCGATGACCCCAAGATACAAATGTCCTGTAAC 2163
Oy 1654 CCGAATACCGGACAGCGCTGTGTTCAACACAGCTGAGACCTGCTGCTCAAGTGTG 1713
Db 2164 CCGAATACCGGACAGCGCTGTGTTCAACACAGCTGAGACCTGCTGCTCAAGTGTG 2223
Oy 1714 GGTTCACGCTTTCCTTAATATGAGGAAAGAGCTGCCACCACTAACAAGAGGCAAGTTC 1773
Db 2224 GGTTCACGCTTTCCTTAATATGAGGAAAGAGCTGCCACCACTAACAAGAGGCAAGTTC 2283
Oy 1774 ATCAGCTTACAGTACCTTCTATTTACCTGATCCGCTTCTGCACTTCAATGCT 1833
Db 2284 ATCAGCTTACAGTACCTTCTATTTACCTGATCCGCTTCTGCACTTCAATGCT 2343
Oy 1834 GCTTACAGCGGGGCTGTGTCACATGTCGACCTTGTGTCATGTCCTCAACCTCTG 1893
Db 2344 GCTTACAGCGGGGCTGTGTCACATGTCGACCTTGTGTCATGTCCTCAACCTCTG 2403
Oy 1894 GCGATACAGCTGTGCTTCTGCTGCGCCCAAGTCTACATGTCCTTCTACCCGGAAGGC 1953
Db 2404 GCGATACAGCTGTGCTTCTGCTGCGCCCAAGTCTACATGTCCTTCTACCCGGAAGGC 2463
Oy 1954 AACAGCGCGCTTCTTCAACAGATGATCCAGGGCTACACCATGAGAGGAGCTAG 2010
Db 2464 AACAGCGCGCTTCTTCAACAGATGATCCAGGGCTACACCATGAGAGGAGCTAG 2520

```

RESULT 6

```

US-10-096-144-3
; Sequence 3, Application US/10096144
; Publication No. US20030022288A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeier, Juergen
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
; FILE REFERENCE: 02307E-088720US
; CURRENT APPLICATION NUMBER: US/10/096,144
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/361,631
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/095,464
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2993
; TYPE: DNA
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: rat G-protein coupled receptor (GPCR) B4
; OTHER INFORMATION: nucleotide sequence
US-10-096-144-3

```

Query Match 61.3%; Score 1232.8; DB 9; Length 2993;

Best Local Similarity 76.9%; Pred. No. 0; Mismatches 442; Indels 30; Gaps 4;

```

Oy 1 ATCACTACAGGCGCATCAGCGATGAGCTGCGAGCAAGTGCTGCCGCTTGTCTG 60
Db 505 ATCACTACAGGCGCATCAGCGATGAGCTGCGAGCAAGTGCTGCCGCTTGTCTG 564
Oy 61 CGTACACACCCAGCGCGGACCCAGCTCCAGGCTTGTGTGCTGATGCTGCACTTC 120
Db 565 CGCAGAGTCCAGCGCGGACCCAGCTCCAGGCTTGTGTGCTGATGCTGCACTTC 624
Oy 121 CGCTGGAATGATGATGCTGTGTGAGCAGCAGCTTGTGCGGCGCAATGAGCCAG 180
Db 625 CAATGGAATGATGATGCTGTGTGAGCAGCAGCTTGTGCGGCGCAATGAGCCAG 684
Oy 181 CTGCTTGGGAGCGCGTGGCC--CGGCGGACATCTGTCATGCTTCCAGGAGCTG 237
Db 685 CTGCTTGGGAGCGCGTGGCC--CGGCGGACATCTGTCATGCTTCCAGGAGCTG 744
Oy 238 CCGACACTGACGCCCAACCAAGATGATGATGAGAGGCGGACGCTGCTGATGACATT 297
Db 745 CCGACACTGACGCCCAACCAAGATGATGATGAGAGGCGGACGCTGCTGATGACATT 804
Oy 298 GTGAGCAAGCTGACAGCAGCAGCAGCGCGCTGTGCTGTGCTTCCGCCCACTGACC 357
Db 805 CTGAGCAAGCTGACAGCAGCAGCAGCGCGCTGTGCTGTGCTTCCGCCCACTGACC 864
Oy 358 CTGTAACACTTCTTCAATGAGTGTGCTGCGCGCAAACTTACAGGCGCGCTGTGATGCGC 417
Db 865 CTGTAACACTTCTTCAATGAGTGTGCTGCGCGCAAACTTACAGGCGCGCTGTGATGCGC 924
Oy 418 TCCGAGTCTGTGGCCATGAGCCGCTGCTGCAACCTACAGGAGCTGCGGCACTTGGGC 477
Db 925 TCCGAGTCTGTGGCCATGAGCCGCTGCTGCAACCTACAGGAGCTGCGGCACTTGGGC 984
Oy 478 ACTTCTCTGGGATCACCATCCAGAGCTGCCATCCCGGGCTTCACTGATGATCCGAGAG 537
Db 985 ACTTCTCTGGGATCACCATCCAGAGGCTGTCCATCCCTGCTTCACTGATGATCCGAGAG 1044
Oy 538 TGGGGCCCAAGGCTGTGGCGCGCACCTTCAAGAGAGCAGGCAAGCTATACCTGCAAC 597
Db 1045 TGGGGCCCAAGGCTGTGGCGCGCACCTTCAAGAGAGCAGGCAAGCTATACCTGCAAC 1104

```


Db 685 CTGTGACCCAGCGCTGACCAAAACGAGCAATCTGATTCCTTCAGAGAGTTCTG 744
Oy 238 CCCAAGCTGACCCCAACCAAGATGAGTGAAGAGCGCCGCTGATGCAAT 297
Db 745 CCCATACCTGATGACGAGGATGAGTGAAGAGCACTGAGCAATC 804
Oy 298 GTGACAGCTGACGAGCAACGAGCGCGCTGAGTGAAGTTCGCGCACTGAC 357
Db 805 CTGACAAAGCTGCGGAGCACTGCGCGCGCTGAGTGAAGTTCGCGCACTGAGC 864
Oy 358 CTGTACCACTTCTTCATGAGTGTGCGCAAGATTCAGCGCGCGCTGAGTGAAGC 417
Db 865 CTGTATACCTTCTTCAGAGAGTGTGCGCAAGATTCAGCGCGCTGAGTGAAGC 924
Oy 418 TCCGAGTCTGCGGCACTGAGCGCGCTGCGCAAGATTCAGCGCGCTGAGTGAAGC 477
Db 925 TCTGAGTCTGCGGCACTGAGCGCGCTGCGCAAGATTCAGCGCGCTGAGTGAAGC 984
Oy 478 ACCTTCTGCGGCACTGAGCGCGCTGCGCAAGATTCAGCGCGCTGAGTGAAGC 537
Db 985 ACTTTCGCGGCACTGAGCGCGCTGCGCAAGATTCAGCGCGCTGAGTGAAGC 1044
Oy 538 TGGGCGCCAGAGCTGCGGCACTGCGCGCGCTGCGCAAGATTCAGCGCGCTGAGC 597
Db 1045 CGCGGTGACAAAGCGAGGTATCCGCTGCTAACAGCAACCTGCGGCACTGAGC 1104
Oy 598 CAGAGTGCAGCACTGCTGAGCGCGCTGCTAACAGCAACCTGCGGCACTGAGC 657
Db 1105 CAGAGTGCAGCGCTGCTGAGCGCGCTGCTAACAGCAACCTGCGGCACTGAGC 1164
Oy 658 GGGGAGCTGCTGAGCGCGCTGAGCGCGCTGAGCGCGCTGAGCGCGCTGAGC 717
Db 1165 GGGGAGCTGCTGAGCGCGCTGAGCGCGCTGAGCGCGCTGAGCGCGCTGAGC 1224
Oy 718 AGCTCTGCGGCTGAGCAAAAGCACTGAGCGCGCTGAGCGCGCTGAGCGCGCTG 777
Db 1225 AGACTCTGCGGCTGAGCAAAAGCACTGAGCGCGCTGAGCGCGCTGAGCGCGCTG 1284
Oy 778 CTGAGGAGATCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 837
Db 1285 CTGAGGAGATCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1344
Oy 838 CAGGAGAGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 897
Db 1345 CAGGAGAGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1404
Oy 898 CCTTTCAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 954
Db 1405 CCTTTCAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1464
Oy 955 ACATCTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1014
Db 1465 GTGTCTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1524
Oy 1015 TCAGGAG 1074
Db 1525 CCAGGAG 1584
Oy 1075 CTTCCGAG 1113
Db 1585 ATGCGAG 1644
Oy 1114 AACGAG 1173
Db 1645 TCCATGAG 1704
Oy 1174 TGGCATGAG 1233
Db 1705 TGGCATGAG 1764
Oy 1234 CTGCGCATCTGAG 1293

Db 1765 CTGCGCATCTGAG 1824
Oy 1294 GCGCCATGAG 1353
Db 1825 GCGCCATGAG 1884
Oy 1354 TACGAG 1413
Db 1885 TATGAG 1944
Oy 1414 TCCATGAG 1473
Db 1945 TCCATGAG 2004
Oy 1474 ATGCGAG 1533
Db 2005 ATGCGAG 2064
Oy 1534 TCTATGAG 1593
Db 2065 TCTATGAG 2124
Oy 1594 CCGAG 1650
Db 2125 ACCATGAG 2184
Oy 1651 AACCCATGAG 1710
Db 2185 CACCCATGAG 2244
Oy 1711 GTGAG 1770
Db 2245 CTGAG 2304
Oy 1771 TTTATGAG 1830
Db 2305 TTTATGAG 2364
Oy 1831 TCTGAG 1890
Db 2365 TCTGAG 2424
Oy 1891 CTGAG 1950
Db 2425 CTGAG 2484
Oy 1951 CGCAAG 2010
Db 2485 CGCAAG 2544

RESULT 8
US-10-124-598-3
Sequence 3, Application US/10124598
Patent No. US20020119526A1
GENERAL INFORMATION:
APPLICANT: Adler, Jon Elliot
APPLICANT: Lindemeyer, Juergen
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
FILE REFERENCE: 02307E-088720US
CURRENT APPLICATION NUMBER: US/10/124,598
CURRENT FILING DATE: 2002-04-16
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/361,631
PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747
PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 10
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 3
LENGTH: 2993

; TYPE: DNA
 ; ORGANISM: Rattus sp.
 ; FEATURE:
 ; OTHER INFORMATION: rat G-protein coupled receptor (GPCR) B4
 ; OTHER INFORMATION: nucleotide sequence
 US-10-124-598-3

Query Match 61.38; Score 1232.8; DB 12; Length 2993;
 Best Local Similarity 76.98; Pred. No. 0;
 Matches 1568; Conservative 0; Mismatches 442; Indels 30; Gaps 4;

QY 1 ATCACTACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 60
 DB 505 ATCACTACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 60
 QY 61 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 120
 DB 565 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 120
 QY 121 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 180
 DB 625 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 180
 QY 181 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 237
 DB 685 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 237
 QY 238 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 297
 DB 745 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 297
 QY 298 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 357
 DB 805 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 357
 QY 358 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 417
 DB 865 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 417
 QY 418 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 477
 DB 925 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 477
 QY 478 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 537
 DB 985 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 537
 QY 538 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 597
 DB 1045 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 597
 QY 598 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 657
 DB 1105 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 657
 QY 658 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 717
 DB 1165 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 717
 QY 718 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 777
 DB 1225 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 777
 QY 778 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 837
 DB 1285 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 837
 QY 838 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 897
 DB 1345 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 897
 QY 898 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 954
 DB 1404 CGTACACAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 954

DB 1405 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1464
 QY 955 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1464
 DB 1465 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1524
 QY 1015 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1074
 DB 1525 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1524
 QY 1075 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1113
 DB 1585 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1584
 QY 1114 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1173
 DB 1645 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1704
 QY 1174 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1233
 DB 1705 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1764
 QY 1234 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1293
 DB 1765 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1824
 QY 1294 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1353
 DB 1825 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1884
 QY 1354 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1413
 DB 1885 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1944
 QY 1414 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1473
 DB 1945 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 2004
 QY 1474 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1533
 DB 2005 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 2064
 QY 1534 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1593
 DB 2065 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 2124
 QY 1594 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1650
 DB 2125 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 2184
 QY 1651 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1710
 DB 2185 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 2244
 QY 1711 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1770
 DB 2245 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 2304
 QY 1771 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1830
 DB 2305 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 2364
 QY 1831 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1890
 DB 2365 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 2424
 QY 1891 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 1950
 DB 2425 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 2484
 QY 1951 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 2010
 DB 2485 CCCTTCCAGAGCCGATGAGTGTGGAACAAGTCCGCTTCCGCTTGTGCTG 2544

```

RESULT 9
US-10-096-144-4
; Sequence 4, Application US/10096144
; Publication No. US20030022288A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeier, Juergen
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
; FILE REFERENCE: 02307E-088720US
; CURRENT APPLICATION NUMBER: US/10/096,144
; CURRENT FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/361,631
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/095,464
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2532
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) B4
US-10-096-144-4

Query Match          60.3%; Score 1212.6; DB 9; Length 2532;
Best Local Similarity 76.3%; Fred. No. 0;
Matches 1555; Conservative 0; Mismatches 454; Indels 30; Gaps 4;

  2 TCACCTACAGCGCCATCATGAGTGTGCGAGACAGAGTGGCTTCCCGCTTGTGCTGC 61
  494 TCACATATAGCCCATCATGAGTGTGCGAGACAGAGTGGCTTCCCGCTTGTGCTGC 553
  62 GTACACACCCAGCGCCGACACACAGTGTGAGGCGCATGGTGTGATGTGCACTTCC 121
  554 GCATGTGTGCGCCGACCCACACACATGTGAGGCGCATGGTGTGATGTGCACTTCC 613
  122 GGTGATGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 181
  614 AGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 673
  182 TGGTGTGCGAGCGGTGCG---CCGCGGACATGTGATCGCTTCCAGAGACGCTGC 238
  674 TGTGTAGCGAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 733
  238 CCACACTGACGCCAACCAGAACATGACGTGAGAGAGCGCGCGCTGTGACCATTTG 298
  734 CGGTGTGAGAACCCAGACGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 793
  298 TGGAAAGTGTGAGAGAGAGAGAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 358
  794 TGGAAAGTGTGAGAGAGAGAGAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 853
  359 TGTACACTTGTTCATATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 418
  854 TGCACAACTTCTTCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 913
  419 CCGAGTCTGTGGCGCATGCAACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 478
  914 CTGAGTCTGTGGCGCATGCAACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 973
  479 CCTTCTGTGGCGCATGCAACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 538
  974 CTTTCTGTGGCGCATGCAACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1033

```

```

  539 GGGGCCACAGCGTGGGCGCCGCCCTGACGAGGACGACGAGCTATACCTGCAAC 598
  1034 GCCATGACAGCCAGGAGTATGCGATGCTTACGAGACAGCTGGGACTTACTTAA 1093
  599 AGGAGTGGACAACTGCTTGAACGCGCCTTGTCTTCAACACATTCCTGAGCTGTG 658
  1094 AGGAGTGGAGCCCTGATGAAACATCAGATCTCTTCAACAGCTTCTGATGCTT 1153
  659 GGGAGCTGTGTGTACACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 718
  1154 GGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1213
  719 GCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1273
  1214 GACTCTTCACTGCAATGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1273
  779 TTAGAGAGATGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 838
  1274 TTAGAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1333
  839 AAGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 898
  1334 AAGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1393
  899 CCTTCAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 955
  1394 CCTTCAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1453
  956 CATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1015
  1454 TGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1513
  1016 CAGGGCAAAAGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1075
  1514 CTGGGCAAAATGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1573
  1076 TTCCCGGCACTTCTTCAACACACTGAA-----TCCCGGAATA 1114
  1574 CGCCGAGACACTTCAACACACTGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGT 1633
  1115 ACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1174
  1634 CATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1693
  1175 GGCAGAGGACCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1234
  1694 GGCAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1753
  1235 TGGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1294
  1754 TGGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1813
  1295 GCCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1354
  1814 GCCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1873
  1355 AGTGGGCGCGCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1414
  1874 ATGTGGGCGCGCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1933
  1415 TCACAAATTTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1474
  1934 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1993
  1475 TGGCAGCGCTTCCAGCGCGCTTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1534
  1994 TGGCAGAGGCTTCCAGCGCGCTTACAGTGTGTGTGTGTGTGTGTGTGTGTGT 2053
  1535 CTATGGCATTTTACAGCGCGCTTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1594
  2054 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2113
  1595 CTCAGTCCACCC---CCGTACTGACCCGAGTACCCCAAGATCAATTTGTCTGTA 1651

```



```

1235 TGGCCATCCTGGTGAATATTCGAGGCACTTCAGACACCATAGTTCGCGGTGGG 1294
1754 TGGCCATCTGCTCATCTTCCTGGAGACATTTCCAGACGCCCATGGTGGCGGGG 1813
1235 GCGCCATGCTCTCTGATGCTGACACTGCTGCTGGTGGCATACATGCTGCCGCT 1354
1814 GCGCCATGCTCTCTGATGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1873
1355 AGCTGGGGGGGGGCAAGGCTCTGACCTGCTGGCGGCAAGGCGCTTTCCTCTGCT 1414
1874 AGTGGGGGGGGGGGCAAGGCTCTGCTGCTGGCGGCAAGGCGCTTTCCTCTGCT 1933
1415 TCACATTTGATCTCTCTGATGCTGACACTGCTGCTGGTGGCATACATGCTGCC 1474
1934 TCTCGCTGCTCTCTGATGCTGACACTGCTGCTGGTGGCATACATGCTGCCGCT 1993
1475 TGGCCAGCGGCTTCCAGCGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGT 1534
1994 TGGCCAGCGGCTTCCAGCGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGT 2053
1535 CTATGGCATTTATACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGT 1594
2054 TGTGGCTTCACTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGT 2113
1595 CTCAGTCCCAACCC---CCGTAAGTACAGGCTTACAGCTGCTGGTGGTGGTGGT 1651
2114 CCATGACCGGCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGG 2173
1652 ACCCACTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1711
2174 ACCCACTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2233
1712 TGGGTTTCACTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 1771
2234 TGGGTTTCACTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 2293
1772 TCATACCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1831
2294 TCATACCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2353
1832 CTGCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1891
2354 CTGCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2413
1892 TGGCATCAGGCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 1951
2414 TGGCATCAGGCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 2473
1952 GCACACGCGGCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 2532
2474 GCACACGCGGCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2532
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) B4
US-10-124-598-4

Query Match      60.3%; Score 1212.6; DB 12; Length 2532;
Best Local Similarity 76.3%; Pred. NO. 0;
Matches 1555; Conservative 0; Mismatches 454; Indels 30; Gaps 4;

1235 TGGCCATCCTGGTGAATATTCGAGGCACTTCAGACACCATAGTTCGCGGTGGG 61
1754 TGGCCATCTGCTCATCTTCCTGGAGACATTTCCAGACGCCCATGGTGGCGGGG 553
1235 GCGCCATGCTCTCTGATGCTGACACTGCTGCTGGTGGCATACATGCTGCCGCT 121
1814 GCGCCATGCTCTCTGATGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 613
1355 AGCTGGGGGGGGGCAAGGCTCTGACCTGCTGGCGGCAAGGCGCTTTCCTCTGCT 181
1874 AGTGGGGGGGGGGGCAAGGCTCTGCTGCTGGCGGCAAGGCGCTTTCCTCTGCT 673
1415 TCACATTTGATCTCTCTGATGCTGACACTGCTGCTGGTGGCATACATGCTGCC 238
1934 TCTCGCTGCTCTCTGATGCTGACACTGCTGCTGGTGGCATACATGCTGCCGCT 733
1475 TGGCCAGCGGCTTCCAGCGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGT 298
1994 TGGCCAGCGGCTTCCAGCGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGT 793
1535 CTATGGCATTTATACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGT 358
2054 TGTGGCTTCACTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 853
1595 CTCAGTCCCAACCC---CCGTAAGTACAGGCTTACAGCTGCTGGTGGTGGTGGT 418
2114 CCATGACCGGCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 913
1652 ACCCACTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 973
2174 ACCCACTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 478
1712 TGGGTTTCACTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 538
2234 TGGGTTTCACTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 598
1772 TCATACCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 658
2294 TCATACCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1153
1832 CTGCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 718
2354 CTGCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1213
1892 TGGCATCAGGCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 778
2414 TGGCATCAGGCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 1273
1952 GCACACGCGGCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 838
2474 GCACACGCGGCTTACAGGCTTACAGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 1333

```


QY 542 GCCCAGAGCTGGCCGCCACCCCTCAGACGAGCAGCCAGAGCTATACCTGC-----A 595
 DB 1075 ATGTACAGGGCTGTAAACACTGTGTCAGCGCTTCCCGAGGGGTCCGTGTCACACTA 1134
 QY 586 ACCAGAGCTGGACAACTGCTGTAAGCCACCTTGTCTTCAACCACTTCTGAGCTCT 655
 DB 1135 ACCAGCTGTGCCGGAGTGCCACACGTCAGACTCTGTAATGCCACAGCTTGGAGCTT 1194
 QY 656 CTGGGAGAGCTGTGCTCAGACGCTGTAAGCTGTGCGGTCTATGCTGTGGCCATGCTCC 715
 DB 1195 TCTCCATGATGTCGCTCAGACAGTGTATGAGGTGTGTAGCTGTGGCTGCGCCAGCTCC 1254
 QY 716 ACAGCTCTCTGGCTGTGTACAAAGACCTGACCAAGAGGTGTACCTTCCCTGGACG 775
 DB 1255 ACCAGCTCTGGAGTGTACTTCTGTGAGATCTGTGTCCAGAGCCCACTTACCTCTGGACG 1314
 QY 776 TGTCTGAGAGATCTGGAAGTCAACTTCTCTCTGAGACCACTTCTTCTTGCAGC 835
 DB 1315 TCTCTCAGCAGATCTACAAAGGTGATTTCTTCTACATGAGAACTCTGTGGCATTTGATG 1374
 QY 836 CGAAGGGGAGCTGTGCTGTGACCTGTGAGATTTGTCACTGTGCAATGGGACCGGAGCAGA 895
 DB 1375 ACAACGGGGACACTCTAGGTACTACGACATCATCTGCTGGAGCTGGAATGGACCTGAT 1434
 QY 896 ATCCCTTCAGAGCTGTGCT 955
 DB 1435 GGACCTTTGATATCATTTGGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1494
 QY 956 CATCTCT-----GCACACCGCTCAACAACAGATCTCTATGTCTATGTCTCAAGAGT 1009
 DB 1495 CAAAATCTAGTGTGACGCGGAAAGAAACATCAGGTCTGTGTGCACTGTGTATGACAGGACT 1554
 QY 1010 GCCAGTCAGGCGCAAAAGAAAGAGTGTGGGATGCCAGTGTGCTGTGCTGTGCTGTGCTGTG 1069
 DB 1555 GTCTGGCAGGACACACAGGT 1614
 QY 1070 ACTGCTTCCCGGACCT 1111
 DB 1615 CCTGGGAAGCTGTGGACCTTCT 1674
 QY 1112 ATAAAGATGTCT 1171
 DB 1675 CAGAAGATGTGGCACCAGGAGAGAGACTACTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1734
 QY 1172 AATGATGAGGACACCACT 1231
 DB 1735 CTGTGCAATGAAACCACT 1794
 QY 1232 CCTGTGCAATCTGTGATATTTCTGTGAGGACCTTCTCTCTCTCTCTCTCTCTCTCTCT 1291
 DB 1795 TGTGTGGAGCTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1854
 QY 1292 GGGGCGCCATGTCT 1351
 DB 1855 GGGGAGGCTGTCT 1914
 QY 1352 TGTACGTGGGCGCCCAAGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1411
 DB 1915 GCTTCTTGGGGAGCCCAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1974
 QY 1412 GCTTCAATTTGATCT 1471
 DB 1975 GGTGTGCAATCT 2034
 QY 1472 AGATGGCCAGCGCTTCT 1531
 DB 2035 AATTTTCTACCAAGGT 2094
 QY 1533 TCTCTATGGCATTTTCAAGGT 1591
 DB 2095 TATTTGCTATTTGTCAAGT 2154
 QY 1592 GGCCTCAGTCCACCCCGCTACTGACCCGATGACCCCAAGATCACAATTTGTCTCTGT 1651

DB 2155 GGAACCCACGACCCACCAAGGGAATATACAGCGCTTCCCATCTGTGATTTCTGAGTGA 2214
 QY 1652 ACCCACTACCGCAACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1711
 DB 2215 CAGAGGTCAACTCTGTAGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2274
 QY 1712 TGGTTTCAAGCTTGTGCTCAATGAGGCAAGAGCTGTGCGCCACCAACATCAAGGCGCA 1771
 DB 2275 GTACTTGTGTGCACTGTGCTGTGAGGAAAGTGTGAGAGTGTGAGAGATTAATGAGCAAT 2334
 QY 1772 TCAATACCTCTGACATGTGATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1831
 DB 2335 GTGTACCTGTAGCGCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2394
 QY 1832 CTGCTTACAGCGGGGT 1891
 DB 2395 GCATTTACAGGGGAGCTACCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2454
 QY 1892 TGGCATCAGCGCTGTGCTGTGCGCCCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1951
 DB 2455 GCGGCGCTTACGCGGT 2514
 QY 1952 GCACACGCGCCGCTACTTCAACAGCATGATCTCAGGGCTACACATGAGAGGGAGC 2007
 DB 2515 TCACAAATACAAACACTTTCAGGCGCTCATCTCAGGAGCTACAGAGGCGCGCTGCGG 2570

RESULT 13
 US-09-927-315-4
 ; Sequence 4, Application US/09927315
 ; Publication No. US20030040045A1

GENERAL INFORMATION:

; APPLICANT: Zuker, Charles S.
 ; APPLICANT: Ryba, Nicholas J.P.
 ; APPLICANT: Nelson, Greg
 ; APPLICANT: Hoon, Mark A.
 ; APPLICANT: Chandrasekar, Jayaram
 ; APPLICANT: Zhang, Yifeng
 ; APPLICANT: The Regents of the University of California
 ; APPLICANT: The Government of the United States of America
 ; APPLICANT: as represented by the Secretary of the
 ; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
 ; FILE REFERENCE: 02307E-120110US
 ; CURRENT APPLICATION NUMBER: US/09/927,315
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: US 60/302,898
 ; PRIOR FILING DATE: 2001-07-03
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 2771
 ; TYPE: DNA
 ; ORGANISM: Rattus sp.
 ; FEATURE:
 ; OTHER INFORMATION: rat T1R1 sweet taste receptor
 US-09-927-315-4

Query Match 19.5%; Score 392.8; DB 9; Length 2771;
 Best Local Similarity 51.4%; Pred. No. 4; 5e-101;
 Matches 1046; Conservative 0; Mismatches 942; Indels 48; Gaps 4;

QY 2 TCACTACAGCGCCATCAAGCATGAGCTGTGAGGAGTGTGGCTGTGGCTGTGGCTGTGGCT 61
 DB 553 TCAGCTATGAGGCAAGCAAGCTGTGTACTGATGCAAGGCAAGTTCCTGCTCTCTCT 612
 QY 62 GTACACACCCAGCGCGGCAACCAAGCTGTGAGGAGGAGTGTGAGTGTGACTGTCTCTCT 121
 DB 613 GTACCGTCCAGGTGACCGGCAACAGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 672
 QY 122 GCTGGAATGTGATCTGT 181

D 673 GGTGGTGTGATCTCGCTATTGGCAGCTACGGTATTACGGGACGCTGGGTGTGACG 732
Q 182 TGTGGGAGAGCGGTGGCCGCGGCGACATCTGCATCGCTTCACAGAGACGCTGCCA 241
D 733 CGCTGAGAGAGCTGGCGGTGGCCGCGGCGATCTGCCTTCAAGGACATCTGCTCC-- 790
Q 242 CACTGACGCCCAACAGACATGACGTACAGAGAGCGGCGGCGCTGGTACATTTGTG 301
D 791 -----TTCCTGCGCGGGGTGGAGCCGAGATGACAGATGATGATG 834
Q 302 ACAGCTGACAGAGACAGACGCGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
D 835 AGCATCTGGCTACAGAGACAGACAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
Q 362 ACCACTTCTTCAATGAGTGTGCTGCGGAGACTTACAGGCGCGGCTGGTGTGCTGCTG 421
D 895 GAGTGTCTTACAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 954
Q 422 AGTCTGGGCGGCTGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481
D 955 AAGACTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014
Q 482 TCTGGGATCATCATCAGAGCGGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541
D 1015 TGTCTGGTGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1074
Q 542 GCCCAGAGCTGGGCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 595
D 1075 ATGTGAGGCGCTGTAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134
Q 596 ACCAGAGGTGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655
D 1135 ACCAGAGGTGAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1194
Q 656 CTGGGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 715
D 1195 TCTCATGAGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1234
Q 716 ACAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 775
D 1255 ACCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1314
Q 776 TGTCTGAGAGATCTGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835
D 1315 TGTCTGAGAGATCTGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1374
Q 836 CGCAGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 895
D 1375 ACAAGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1434
Q 896 ATCCCTTCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955
D 1435 GAGCTCTTGAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1494
Q 956 CATCTCT-----GCACAGCTCAACACAGATCCATATGCTGCTGCTGCTGCTGCT 1009
D 1495 CAAATTCAGAGTGGCGGAGGAGAGATCAATGAGTGTGCTGCTGCTGCTGCTGCTGCT 1554
Q 1010 GCCAGTACGAGCAAAAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1069
D 1555 GTCTGGAGGAGGAGCAACAGAGGTGTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1614
Q 1070 ACTGCTCTCCGCGACCTCTCTCAACACATGAA-----TGCCGGA 1111
D 1615 CCTGCGAAGCTGGAGCTTCTCTCAACATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1674
Q 1112 ATACAGTGTGCTTACACAGTACAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1171
D 1675 CAGAGAGATGGGAGCCCAAGAGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1734
Q 1172 AATGAGTACAGGACACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1231
D 1735 CTGCGCATGAAACCATCTCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1794

Q 1232 CCTGGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1291
D 1795 TGTGGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1854
Q 1292 GGGGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1351
D 1855 GGGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1914
Q 1352 TGTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1411
D 1915 GCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1974
Q 1412 GCTTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1471
D 1975 GCTTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2034
Q 1472 AGATGGCAGCGCTTCCACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1531
D 2035 AGTCTTACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2094
Q 1532 TCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1591
D 2095 TATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2154
Q 1592 GGCCTGAGTCCACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1651
D 2155 GAGCCCGACGAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2214
Q 1652 ACCCCATACCGGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1711
D 2215 CAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2274
Q 1712 TGGGTTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1771
D 2275 GTACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2334
Q 1772 TCATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1831
D 2335 GTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2394
Q 1832 CTGCTACAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1891
D 2395 GCATTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2454
Q 1892 TGGCATACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1951
D 2455 GCGGCGGCTTACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2514
Q 1952 GCAACAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2007
D 2515 TCACATATACGAGACATTTTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2570

RESULT 14
US-09-819-946-3
Sequence 3, Application US/09819946
Patent No. US2002045740A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Scoville, John
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Abidin, Alejandro
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. US2002045740A1 Human 7TM Proteins and Polynucleotides E
FILE REFERENCE: LEX-0157-USA
CURRENT APPLICATION NUMBER: US/09/819,946
CURRENT FILING DATE: 2001-08-27
PRIOR APPLICATION NUMBER: US 60/192,978
PRIOR FILING DATE: 2000-03-28

NUMBER OF SEQ ID NOS: 9
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 2292
 TYPE: DNA
 ORGANISM: homo sapiens
 US-09-819-946-3

Query Match 19.2%; Score 386.4; DB 10; Length 2292;
 Best Local Similarity 51.1%; Pred. No. 2.7e-99;
 Matches 1042; Conservative 1; Mismatches 947; Indels 48; Gaps 4;

1 ATCACTTACAGGCGCATGATGAGTGGCGAGACAAAGGTGCGCTCCGCTTGGCG 60
 265 ATTACCTTGGGGCGGACGAGCAGACGCTCAGCGGTGAAGCGGCAATGCCCTTTCTG 324
 61 GGTACCAACCCAGCGCGGACCAACAGTGCAGGCGCATGTGTGATGCTGACATTC 120
 325 GGCACATCCCAATGAGAAATGACAGGTGAGACCATGTGCTGCTGCGCAAGATTC 384
 121 CGGTGAGTGGATGATGTTGTGTGTGAGCAGCGACCATGAGCGCGGACAAATGGCCAG 180
 385 GGTGAGACTGATCTCTGTGTGTGTGAGCAGCTGACATGAGCAGCTAGGGGTGCG 444
 181 CTGCTTGGCGAGCGCGTGGCGCGGCGACATCTGATCGCTTCCAGAGAGCGTGC 240
 445 GCACCTGGAGAACCCAGGCGCACTGTGAGGGGATCTGATTCATTTCAAGGACATCAATGCC 504
 241 ACACGTGACCCCAACCAACATGACGTGACAGAGAGCGCGCTGTGACCAATTTG 300
 505 -----TTCTGTGCGGAGGTGGCGATGAGAGATGAGATGCGCTCATG 546
 301 GACAAGCTGACAGACAGACAGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
 547 CGCCACCTGGCGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 606
 361 TACCACTTCTCAATGAGGT 420
 607 AGGGT 666
 421 GAGTCTTGGCGATGACCGCGTCTGTGACAACTGACAGAGAGTGGCGCATTTGGCGACC 480
 667 GAGGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 726
 481 TTCTGCGGATCACCATCAGAGCGTGCACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
 727 GTGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 786
 541 GCGCCACAGGCTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 594
 787 TATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 846
 595 AACCAAGAGTGGGACATCGCTGAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 654
 847 AATCAGCTCTGAGAAATGCGCAAGCTTTCATGRCACACAGCATGCCCAAGTCAAAAGCG 906
 655 TCTGGGAGCGGT 714
 907 TTCTGCAATGATTTGCTGT 966
 715 CACAGCTCTGT 774
 967 CACAGCTCTGT 1026
 775 CTGCTTGGAGAGATGGAAGGTCAACTTCACTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 834
 1027 CTTTGGAGAGATCACAAGGTGCAATTTCTTTCACACAGGACATGCTGCGCTTTAAT 1086
 835 CGCAAGGAGAGT 894
 1087 GACAACAGAGATCCCTGATAGTATAAATTAATTTGCTGTGTGTGTGTGTGTGTGTGTGTGT 1146
 895 AATCCCTTCCAGAGGCTGCTCTAC-----TACCCCTGACAGGACAGTGAAGAAC 948

1147 TGGACCTTCAGGCTGT 1206
 949 ATCAAGACATCTGTGACACCGGTCAACAAACAGATCCATGTGTGTGTGTGTGTGTGTGTGTGT 1008
 1207 ACCAAATTCAGTGTGACGAGAAAGCAACAGGTGCTAATGTGTGTGTGTGTGTGTGTGTGTGT 1266
 1009 TGGCAGTGGGCAAAAGAGAGAGGT 1068
 1267 TGTCTTGAAGGCGACAGGAGT 1326
 1069 GACTGCTTCCGCGGACCTTCCCTCAACACACTGAA-----TGCCTG 1110
 1327 CCTGTGGGCGTGGGAGCTTCTCAACAAAGATGACCTGTACAGATGCCAGGCTGTGGG 1386
 1111 AATAAGAGT 1170
 1387 AAAGAGAGT 1446
 1171 GAATGGCATGAGGCAACCCACATCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1230
 1447 GCTTGTGCGACACACCTTGT 1506
 1231 ACCCTGGCCATCTGT 1290
 1507 CTGCTTGGGAGT 1566
 1291 GGGGCGCCCATGT 1350
 1567 GGGGCGCGCTGT 1626
 1351 GTGTAGTGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1410
 1627 GCGTCTTGTGGGGAACCCAAAGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1686
 1411 TGTCTACATTTGATCTGT 1470
 1687 GGTTCACCATCTTCTGT 1746
 1471 AAGATGGCGAGCGGCTTCCACAGCGGCTGACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1530
 1747 AAGTTCCTCACAAGGTACATCTGATCTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1806
 1531 GTCTATAGGCAATTTATCAGGCTACTCAAAATGTCATTTGTGTGTGTGTGTGTGTGTGTGTGT 1590
 1807 CTGT 1866
 1591 CGGCTCAGTCCACCGCGT 1650
 1867 TGGACCCCACTGT 1926
 1651 AACCCCACTACGGAACAGCGGCTGT 1710
 1927 ACAGAGACCAATCTCTGT 1986
 1711 GTGGGTTTCAGGCTGT 1770
 1987 AGTGCTTGT 2046
 1771 TTTCATCAGCTTGT 1830
 2047 TGTGTACCTTGT 2106
 1831 TGT 1890
 2107 AGCGTCTGT 2166
 1891 CTGGCATCAGCTGT 1950
 2167 ACAGAGGCTGT 2226
 1951 CGCAACGCGCGCTACTTCAACAGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2008

Db 2227 CTCACAGCAGACAGACCTTCAGAGCTTCCATTCAGAGACTACAGAGCGCTCGGCT 2284

RESULT 15

US-09-819-946-1

Sequence 1, Application US/09819946

Patent No. US20020045740A1

GENERAL INFORMATION:

APPLICANT: Walke, D. Wade

APPLICANT: Scoville, John

APPLICANT: Donoho, Gregory

APPLICANT: Turner, C. Alexander Jr.

APPLICANT: Friedrich, Glenn

APPLICANT: Abidin, Alejandro

APPLICANT: Zambrowicz, Brian

APPLICANT: Sands, Arthur T.

TITLE OF INVENTION: No. US20020045740A1 Human 77M Proteins and Polynucleotides Encl

FILE REFERENCE: LEX-0157-USA

CURRENT APPLICATION NUMBER: US/09/819,946

CURRENT FILING DATE: 2001-08-27

PRIOR APPLICATION NUMBER: US 60/192,978

PRIOR FILING DATE: 2000-03-28

NUMBER OF SEQ ID NOS: 9

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 1

LENGTH: 2526

TYPE: DNA

ORGANISM: homo sapiens

US-09-819-946-1

Query Match 19.2%; Score 386.4; DB 10; Length 2526;

Best Local Similarity 51.1%; Pred. No. 2.8e-99;

Matches 1042; Conservative 1; Mismatches 947; Indels 48; Gaps 4;

QY 1 ATCACTCAGACGCGCATCAGAGTACGAGTGCAGACAGAGGCGCTCCCGGGCTTGTCTG 60

Db 499 ATTGATGATGGCGCCAGAGGAGAGGCTCAGCTGAAGCGGATATCCCTTTCTCTG 558

QY 61 CGTACACACCGCAGCGCGCAGACACAGTGCAGGCGCATGTGTCAGCTGATCTGCATTC 120

Db 559 CGCAGCATCCCATGACATGACATGACATGACATGACATGACATGACATGACATGACATG 618

QY 121 CGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180

Db 619 GGGGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 678

QY 181 CTGCTGCGAGCGCGTGGCGCGCGCGAGATTCGATTCGCTCCAGGAGAGCTGCGC 240

Db 679 GCACTGAGAGACAGGCGCATCTGTCAGGGATTCGATTCGATTCGATTCGATTCGATTC 738

QY 241 ACACTGACGCCCAACAGACATGATGATGATGATGATGATGATGATGATGATGATG 300

Db 739 -----TTCTGTGCGCCAGGTGGGCGATGAGAGATGAGATGATGATGATGATG 780

QY 301 GACAGCTGCAGCAGACAGCAGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

Db 781 CGCAGCCTGCG 840

QY 361 TACACTTCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420

Db 841 AGGCTGTTTTTTCAGTCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900

QY 421 GAGTCTGGGCGCATGACAGCGCGGTCTGACACAACTCAGCGAGCTGGGCGCATGTTGGG 480

Db 901 GAGGCTGGGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960

QY 481 TTCCTGGGCGCATGACAGCGCGGTCTGACACAACTCAGCGAGCTGGGCGCATGTTGGG 540

Db 961 GTGCTGGGCGCTGCGCATGACAGCGCGGTCTGACACAACTCAGCGAGCTGGGCGCATG 1020

QY 541 GGGCCACAGGCTGGGCG 1080

Db 1021 TATGCG 1080

QY 595 AACGAGAGTGCAGACATGCTGCTGAAGCCACTTGTCTTCAACACCATTCCTGAGGCTC 654

Db 1081 AATACGCTCTGCAGAGATGCAAGCTTTCATGACACAGATGCCCAAGCTCAAGGCC 1140

QY 655 TCTGGGAGCGCTGCTGCTACAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 714

Db 1141 TTTCTCATGATTTCTGCTTACACAGCATACCGGCGCTGATGCGGCGCATGAGCGCTC 1200

QY 715 CACAGCTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 774

Db 1201 CACAGCTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260

QY 775 TCTGCTGAGAGATCTGGAAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834

Db 1261 CTTTGGAGCAGATTCACAAAGTGCATTTCTTCTACACAGACACTGTGCGCTTAT 1320

QY 835 CCGCAGGGGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894

Db 1321 GACACAGAGATCCCTCAGTATACATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380

QY 895 AATCCCTTCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948

Db 1381 TGAACCTTCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440

QY 949 ATCAAGACATCTCTGACACCGCTGACACAGATCCCTATGCTGCTGCTGCTGCTGCTGCTG 1008

Db 1441 ACCAAATCCAGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500

QY 1009 TGCCAGTACAGGCAAAAGAGAGAGCTGTGGGATCCAGCTGCTGCTGCTGCTGCTGCTGCTG 1068

Db 1501 TGTCTTGAAGGCGACAGCGAGTGTGATGAGGCTTTCATCTGCTGCTGCTGCTGCTGCTGCTG 1560

QY 1069 GACTGCTTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1120

Db 1561 CCGTGTGGGCTGGAGCTTCTCAGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620

QY 1111 AATACAGAGTGTCTCTACAGAGTGAACCTCTGCTTCAAGCGCAGCTGCTGCTGCTGCTG 1170

Db 1621 AATACAGAGTGTCTCTACAGAGTGAACCTCTGCTTCAAGCGCAGCTGCTGCTGCTGCTG 1680

QY 1171 GAATGCTAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1230

Db 1681 GCTTGTGGTGAACACACCTTGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740

QY 1231 ACCCTGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1290

Db 1741 CTGCTTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800

QY 1291 GGGGCGCCCATGCTTCTGATGCTGACACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1350

Db 1801 GGGGCGCCCATGCTTCTGATGCTGACACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860

QY 1351 GTTACGTGGGCGCGCCCAAGTCTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1410

Db 1861 GCGTCTTGTGGGAGAACCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920

QY 1411 TGTCTTCAATTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1470

Db 1921 GGTTCACACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980

QY 1471 AAGATGCGCAGCGCTTCCACAGCGGCTTACAGTACTGCTGCTGCTGCTGCTGCTGCTGCTG 1530

Db 1981 AAGTTCACACAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040

QY 1531 GTCTTATGCGATTTATCAGAGTACTCAAAATGCTATGCTGCTGCTGCTGCTGCTGCTGCTG 1590

Db 2041 CTGTTTGTGATGATCAGCTGAGCGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100

QY 1591 CGGCTTATGCTGAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1650

Db 2101 TGAACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 11:21:44 ; Search time 6591.01 Seconds

(without alignments)
11299.347 Million cell updates/sec

Title: US-09-927-315-14

Perfect score: 2559
Sequence: 1 atgtcggagccctgtctcct.....atcagggaacatgagtgta 2559

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_hcg_hum: *
31: em_hcg_inv: *
32: em_hcg_other: *
33: em_hcg_mus: *
34: em_hcg_pln: *
35: em_hcg_rod: *
36: em_hcg_mam: *
37: em_hcg_vrt: *
38: em_sy: *
39: em_hgo_hum: *
40: em_hgo_mus: *
41: em_hgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2557.4	99.9	3458	6	AX417279
2	2555.8	99.9	2559	6	AX239649
3	2555.8	99.9	2559	6	AX298370
4	2555.8	99.9	3489	6	AX239647
5	2514.4	98.3	2729	6	AX354026
6	2090.6	81.7	2739	6	AX318243
7	1949	76.2	4256	9	AB065647
8	1949	76.2	8445	9	AL139287
9	1947.4	76.1	8001	6	AX417281
10	1923.8	75.2	197748	2	AC026283
11	1574	61.5	1788	6	AX451560
12	1470.8	57.5	2577	10	AF456324
13	1470.8	57.5	2577	10	AY032620
14	1454.8	56.9	2577	6	AX298367
15	1454.8	56.9	2577	10	AF311386
16	1454.8	56.9	3521	10	AF337039
17	1454.8	56.9	3554	10	AB049994
18	1450	56.7	2577	10	AY032621
19	1083	42.3	1113	6	AX451557
20	873.2	34.1	3150	10	AY026318
21	871.6	34.1	3153	10	AF368024
22	871.6	34.1	4105	10	AB055708
23	871.6	34.1	11809	6	AX298368
24	871.6	34.1	232941	2	AL670236
25	871.6	34.1	254197	2	AF389853
26	870	34.0	3154	10	AF368025
27	621	24.3	190537	2	AC126156
28	366	14.3	2526	6	AX282892
29	366	14.3	2951	6	AX282900
30	365.6	14.3	2759	6	AX354027
31	359	14.0	2516	9	AF387618
32	357	14.0	2292	6	AX282894
33	354	13.8	2526	10	AF301162
34	350.8	13.7	2520	10	AF127389
35	346	13.5	2892	10	AF337040
36	344.4	13.5	2529	10	AY032622
37	341.2	13.3	2526	10	AF301161
38	328.4	12.8	190537	2	AC126156
39	317.2	12.4	98248	2	AC095793
40	296.8	11.6	3060	10	AF337041
41	284	11.1	2532	6	AR208616
42	284	11.1	2532	10	AY032623
43	279.4	10.9	2529	10	AF127390
44	279.4	10.9	2993	6	AR208615
45	218.6	8.5	2010	6	AR208617

ALIGNMENTS

RESULT 1
AX417279
LOCUS AX417279 3458 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 1 from Patent WO0230981.
ACCESSION AX417279
VERSION AX417279.1 GI:21522589
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Difrancesco V., Beasley E.M., Ketchum K.A., Wei M.-H. and Zhong W.
TITLE Gene encoding a human g-protein coupled receptor and its use
JOURNAL Patent: WO 0230981-A I 18-APR-2002;

FEATURES PE CORP NY (US)
Location/Qualifiers
source 1..3458
/organism="Homo sapiens"
/db-xref="taxon:9606"
BASE COUNT 590 a 1179 c 1092 g 597 t
ORIGIN

Query Match 99.9%; Score 2557.4; Db 6; Length 3458;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2558; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTGGGCGCTGCTGTCTCTGAGGCTTCAAGCTTGGGCTCTCTGACACCTTGGAGCGGG 60
Db 30 ATGCTGGGCGCTGCTGTCTCTGAGGCTTCAAGCTTGGGCTCTCTGACACCTTGGAGCGGG 89

QY 61 GCGCCATTGTGCTGTCTCTGAGGCTTCAAGCTTGGGCTCTCTGACACCTTGGAGCGGG 120
Db 90 GCGCCATTGTGCTGTCTCTGAGGCTTCAAGCTTGGGCTCTCTGACACCTTGGAGCGGG 149

QY 121 TTCCCGCTGGGCGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
Db 150 TTCCCGCTGGGCGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 209

QY 181 GTGTGACACGAGGCTCTCTCTCAAGGCGCTGCTGTGGGCACTGGCCATGAAATGGCGTG 240
Db 210 GTGTGACACGAGGCTCTCTCTCAAGGCGCTGCTGTGGGCACTGGCCATGAAATGGCGTG 269

QY 241 GAGAGATCAACAACAGTGGATCTGTGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db 270 GAGAGATCAACAACAGTGGATCTGTGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 329

QY 301 GATACGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 330 GATACGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 389

QY 361 GCGAGCGCGGACATGCGCGGCTACTGCACTACACGAGTACACGAGGAGGAGGAGGAGG 420
Db 390 GCGAGCGCGGACATGCGCGGCTACTGCACTACACGAGTACACGAGGAGGAGGAGGAGG 449

QY 421 GTCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
Db 450 GTCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 509

QY 481 CTATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
Db 510 CTATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 569

QY 541 CCTCTCTTCTTCCGACCGGCTGCGGAGCGAGCGAGTGTGCACTGACGCGCGCGGAGCTG 600
Db 570 CCTCTCTTCTTCCGACCGGCTGCGGAGCGAGCGAGTGTGCACTGACGCGCGCGGAGCTG 629

QY 601 CTGAGAGAGTTCGGCTGGAAGTGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
Db 630 CTGAGAGAGTTCGGCTGGAAGTGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 689

QY 661 CAGGCGCTGAGCATCTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Db 690 CAGGCGCTGAGCATCTTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 749

QY 721 GGCTGTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
Db 750 GGCTGTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 809

QY 781 CACCAAGTGAACAG 840
Db 810 CACCAAGTGAACAG 869

QY 841 CAGGCGCTTCTCAACTACAGATCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 870 CAGGCGCTTCTCAACTACAGATCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 929

QY 901 GAGGCGTGGCTGACCTGTGACCTGTGATGGGCTGCCCGGAGTGGCCAGAGTGGGAGC 960

Db 930 GAGGCGTGGCTGACCTGTGACCTGTGATGGGCTGCCCGGAGTGGCCAGATGGGAGC 989

QY 961 GTGCTGGCTTCTTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020

Db 990 GTGCTGGCTTCTTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1049

QY 1021 CACCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080

Db 1050 CACCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1109

QY 1081 CTGAG 1140

Db 1110 CTGAG 1169

QY 1141 GTGAGCGGAG 1200

Db 1170 GTGAGCGGAG 1229

QY 1201 GCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260

Db 1230 GCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1289

QY 1261 GTGAGCGGAG 1320

Db 1290 GTGAGCGGAG 1349

QY 1321 CCGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380

Db 1350 CCGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1409

QY 1381 TGGCAGGAGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440

Db 1410 TGGCAGGAGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1469

QY 1441 GAGGCGCTTGAAGATCCGCTGGGAGAGAGTCTGACACAGAGAGAGAGAGAGAGAGAG 1500

Db 1470 GAGGCGCTTGAAGATCCGCTGGGAGAGAGTCTGACACAGAGAGAGAGAGAGAGAG 1529

QY 1501 CCGCAGTGGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560

Db 1530 CCGCAGTGGCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1589

QY 1561 TGTGTGAGATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620

Db 1590 TGTGTGAGATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1649

QY 1621 TGTGTGAGATGCGAGTGTGCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680

Db 1650 TGTGTGAGATGCGAGTGTGCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1709

QY 1681 TTCTTGGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740

Db 1710 TTCTTGGCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1769

QY 1741 GCGCTTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800

Db 1770 GCGCTTGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1829

QY 1801 GCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860

Db 1830 GCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1889

QY 1861 GTCTCTCTTCTTCCGACCGGCTGCGGAGCGAGCGAGTGTGCACTGACGCGCGGAGCTG 1920

Db 1890 GTCTCTCTTCTTCCGACCGGCTGCGGAGCGAGCGAGTGTGCACTGACGCGCGGAGCTG 1949

QY 1921 CACCTCCGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980

Db 1950 CACCTCCGCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2009

QY 1981 GAGTCAGAACTGCTCTGAGCTGGGAGAGCGGCTGAGTGGCTGCTGCGGAGGAGGAGC 2040

[illegible][illegible]

Query Match	Similarity	99.9%	Score	2555.8	DB	6	Length	2559
Best Local	Similarity	99.98%	Pred.	No. 0				
Matches	2557	Conservative	0	Mismatches	2	Indels	0	Gaps
Qy	1	ATGCTGGGCGCTGCTGCTCTGGGCGCTCAGCCCTCTCTGCTCTCTCACCCTGGAGGGG	60					
Db	1	ATGCTGGGCGCGTGTCTCTGGGCGCTCAGCCCTCTCTGCTCTCTCACCCTGGAGGGG	60					
Qy	61	GCCCATGTGCTGCTCTCAGCAAGCACTTATGATGAAGGGGAGCTACGTGCTGGGGGGCTG	120					

D	b	61	gcccacatttgctgctgtacacagcaactttagatgaagaaagggagactacgtgtctggggggctg	120
Q	y	121	ttccccctggggcgagggcgagagagcttgacctccgacgcccagacaggccagagccct	180
D	b	121	ttccccctggggcgagggcgagagagcttgacctccgacgcccagacaggccagagccct	180
Q	y	181	gtgtgcacacaggtttctcttaaacggccctgtcttgggcacttgccattgaaataatggccgtg	240
D	b	181	gtgtgcacacaggtttctcttaaacggccctgtcttgggcacttgccattgaaataatggccgtg	240
Q	y	241	gagagatcaaacaaagaatgsgatgtgctgccccgggctgcggctgggcttagacctcttt	300
D	b	241	gagagatcaaacaaagaatgsgatgtgctgccccgggctgcggctgggcttagacctcttt	300
Q	y	301	gatatgctgtcggagacctgtgtggtgacatgaagaccacgcttcattgttctggccaaagca	360
D	b	301	gatatgctgtcggagacctgtgtggtgacatgaagaccacgcttcattgttctggccaaagca	360
Q	y	361	ggcagccggccacatcgccgcttacttcacactacacgcaatcacaccccggtgtgctgct	420
D	b	361	ggcagccggccacatcgccgcttacttcacactacacgcaatcacaccccggtgtgctgct	420
Q	y	421	gtcatcgggggcccaactcgtcagagctcgccacatggtcacggcgaatttttagcttttc	480
D	b	421	gtcatcgggggcccaactcgtcagagctcgccacatggtcacggcgaatttttagcttttc	480
Q	y	481	ctcatgccccacaggtcagctacggtgtgtacatgatgagactgtgagcgccgggagaccttc	540
D	b	481	ctcatgccccacaggtcagctacggtgtgtacatgatgagactgtgagcgccgggagaccttc	540
Q	y	541	cccttccttttccggcaccggtgcccacgacacgctgtgtcagactgacggccggcgagctg	600
D	b	541	cccttccttttccggcaccggtgcccacgacacgctgtgtcagactgacggccggcgagctg	600
Q	y	601	ctcgagagagttcggtcgatgaactgggtggccgcccctgtggagagagagagatagcgccg	660
D	b	601	ctcgagagagttcggtcgatgaactgggtggccgcccctgtggagagagagagatagcgccg	660
Q	y	661	cagggccttgagacatcttctgcggccctggcccgggcagcggcactgtgcctgcgcacag	720
D	b	661	cagggccttgagacatcttctgcggccctggcccgggcagcggcactgtgcctgcgcacag	720
Q	y	721	ggcctggtggccgctgccccgtggccgctgacatgactggcgctggggamaagtcacagactctg	780
D	b	721	ggcctggtggccgctgccccgtggccgctgacatgactggcgctggggamaagtcacagactctg	780
Q	y	781	caccaggtgaacacagacacgctgacagagtgatgctctttggcctcgacagccgc	840
D	b	781	caccaggtgaacacagacacgctgacagagtgatgctctttggcctcgacagccgc	840
Q	y	841	cacgccccttccaactacacacatcacagacagagcttgcgcccaagagtggtgggtggccac	900
D	b	841	cacgccccttccaactacacacatcacagacagagcttgcgcccaagagtggtgggtggccac	900
Q	y	901	gagggcttgctgaccttctgacctgtgctcattggggcttgccggcattggccacagtggcgac	960
D	b	901	gagggcttgctgaccttctgacctgtgctcattggggcttgccggcattggccacagtggcgac	960
Q	y	961	gtgcttggtgcttccccaagaggggtggccagagtcacagatgtccccaatgagtgaaagc	1020
D	b	961	gtgcttggtgcttccccaagaggggtggccagagtcacagatgtccccaatgagtgaaagc	1020
Q	y	1021	caccttgaccttgagcacacgacccggccttctgcttgacctggggagagggagcagaggt	1080
D	b	1021	caccttgaccttgagcacacgacccggccttctgcttgacctggggagagggagcagaggt	1080
Q	y	1081	ctggagagagagcgtgtggggcgacccggccttctgtctggccctggggagaaaggaagcaggt	1140
D	b	1081	ctggagagagagcgtgtggggcgacccggccttctgtctggccctggggagaaaggaagcaggt	1140
Q	y	1141	gtgagcccgaggcttaaatcaccacagagcttctgtctatggcagctgtgtatagctg	1200
D	b	1141	gtgagcccgaggcttaaatcaccacagagcttctgtctatggcagctgtgtatagctg	1200

OY	1201	GCGAGGCGCCGGCACAACACTCTTCAGTGTGAACGCCCTCAAGCGTGCGCCCGGAGAACCCC	1280
Dd	1201	GCCAAGCGCCCTGGCACAACTCTTCATGTGAAGCCCTCAGGCTGCCCGCGAGAACCCC	1260
OY	1261	GTYAAGCCCTTGAGAGCTCTGTGAGAATCATGTACAACTCACTTCCACGTGGAGGGGCTG	1320
Dd	1261	GTMAGAGCCCTTGAGAGCTCTGTGAGAAATCATGTACAACTCACTTCCACGTGGAGGGGCTG	1320
OY	1321	CCGCTGCGGTTTCGACAGACAGCGGAACGTGAGCATGGAGTACAGACCTGMACTGTGGGTG	1380
Dd	1321	CCGCTGCGGTTTCGACACACACCGGAACGTGAGCATGGAGTACAGACCTGMACTGTGGGTG	1380
OY	1381	TGGCAGGGCTTAGTGCCCAAGCTCCAGAGTGAAGTGGGAGGTTCACAGGCGCCTCAGAGCA	1440
Dd	1381	TGGCAGGGCTTAGTGCCCAAGCTCCAGAGTGAAGTGGGAGGTTCACAGGCGCCTCAGAGCA	1440
OY	1441	GAGCGCCTGAAGATCCGCTGGCACAACGCTGACAAACAGAACCCGCTGCCGGTGTCTG	1500
Dd	1441	GAGCGCCTGAAGATCCGCTGGCACAACGCTGACAAACAGAACCCGCTGCCGGTGTCTG	1500
OY	1501	CGGAGATGCCAGAGGGGCCAGAGTGGCGCGGCTCAAGGGGTTCCACTCCGCTGCTAGACAC	1560
Dd	1501	CGGAGATGCCAGAGGGGCCAGAGTGGCGCGGCTCAAGGGGTTCCACTCCGCTGCTAGACAC	1560
OY	1561	TGTGTGACTTCGAGAGGGGAGCAGCTACCCGCAAAAACCAGAGACATCGCTGCACTTT	1620
Dd	1561	TGTGTGACTTCGAGAGGGGAGCAGCTACCCGCAAAAACCAGAGACATCGCTGCACTTT	1620
OY	1621	TGTGTGCCAGGATTAAGTGTGTCCCCGGAGACCAACACGCTCTCCCGCGCAGGTCTGG	1680
Dd	1621	TGTGTGCCAGGATTAAGTGTGTCCCCGGAGACCAACGCTCTCCCGCGCAGGTCTGG	1680
OY	1681	TTTCCGTGGCATGGGGCGAGACCGGCTATGCTACTCTCTCTCTGCTGTGAGCCTGGCGCTG	1740
Dd	1681	TTTCCGTGGCATGGGGCGAGCCGGCTATGCTACTCTCTCTCTGCTGTGAGCCTGGCGCTG	1740
OY	1741	GGCCTTGTGCTGTGCTTTTGGGGCTGTTCGTTCAACATTGSGAGACGCCACTGGTTCAG	1800
Dd	1741	GGCCTTGTGCTGTGCTTTTGGGGCTGTTCGTTCAACATTGSGAGACGCCACTGGTTCAG	1800
OY	1801	GCGTGGGGGGGGGCCCGGGGCTGTTTTGGCGCTGTGTGGCGTGGGGCTGGTCTGCCCTCAGC	1860
Dd	1801	GCGTGGGGGGGGGCCCGGGGCTGTTTTGGCGCTGTGTGGCGTGGGGCTGGTCTGCCCTCAGC	1860
OY	1861	GTCCTCTGTTCCTCTGGGCGAACGCCAGCCCTGCCGATGCTGTGGCCAGCAGCCCTTGTCC	1920
Dd	1861	GTCCTCTGTTCCTCTGGGCGAACGCCAGCCCTGCCGATGCTGTGGCCAGCAGCCCTTGTCC	1920
OY	1921	CACCCCGCGTACAGGGGCTCCGTGAGACACTTTCCTGTGAGCGCGGAGNATCTTCGTG	1980
Dd	1921	CACCCCGCGTACAGGGGCTCCGTGAGACACTTTCCTGTGAGCGCGGAGNATCTTCGTG	1980
OY	1981	GAGTCAGAACTGCTCTGAGCTGGAGCAGACCGGCTATGAGGCTGTGCTGTGGGGGCGCTGAG	2040
Dd	1981	GAGTCAGAACTGCTCTGAGCTGGAGCAGACCGGCTATGAGGCTGTGCTGTGGGGGCGCTGAG	2040
OY	2041	GCGTGGCTGT	2100
Dd	2041	GCGTGGCTGT	2100
OY	2101	GTTGGCTTCCCGCGAGAGT	2160
Dd	2101	GTTGGCTTCCCGCGAGAGT	2160
OY	2161	CACGTGCGCACAGCTCTGTGGGTACGCTTTCGCGCTTAGCGCACGCCAATGCCACGCTG	2220
Dd	2161	CACGTGCGCACAGCTCTGTGGGTACGCTTTCGCGCTTAGCGCACGCCAATGCCACGCTG	2220
OY	2221	GCGCTTCTGTGCTCTCTGTGGGCACTTTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2280
Dd	2221	GCGCTTCTGTGCTCTCTGTGGGCACTTTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2280

QY	2281	GGCGTGGCGTCACCGTTGGCCATGCTGGCGTCATTCATCAACCGTGGGCTCCCTTGTGGCC	2340
QY	2281	GGCGTGGCGTCACCGTTGGCCATGCTGGCGTCATTCATCAACCGTGGGCTCCCTTGTGGCC <td>2340</td>	2340
QY	2341	CTCCCTGGCCAAATGTCAGAGTGGTCTCTCAAGGCCGCCGCTGCAGATGGGCGCCCTCTGTGTC <td>2400</td>	2400
DB	2341	CTCCCTGGCCAAATGTCAGAGTGGTCTCTCAAGGCCGCCGCTGCAGATGGGCGCCCTCTGTGTC <td>2400</td>	2400
QY	2401	TGTGTCTGGGCGATCTGTGGCTGCTTCCACCTGCCCCAGGTGTTACCTGCTCATGCGCGAG <td>2460</td>	2460
DB	2401	TGTGTCTGGGCGATCTGTGGCTGCTTCCACCTGCCCCAGGTGTTACCTGCTCATGCGCGAG <td>2460</td>	2460
QY	2461	CCAGGGGCTCAACACCCCCAGAGTTCTCCGCGGAGGGGGCCCGGGGATGCCCAAGGCCG <td>2520</td>	2520
DB	2461	CCAGGGGCTCAACACCCCCAGAGTTCTCCGCGGAGGGGGCCCGGGGATGCCCAAGGCCG <td>2520</td>	2520
QY	2521	AATGACGGGAACACAGAAATCATCGGGGGAATCATGATGA <td>2559</td>	2559
DB	2521	AATGACGGGAACACAGAAATCATCGGGGGAATCATGATGA <td>2559</td>	2559
RESULT 3			
LOCUS	AX298370	2559 bp	DNA linear PAT 26-NOV-2001
DEFINITION	Sequence 4 from Patent WO0183749.		
ACCESSION	AX298370		
VERSION	AX298370.1	GI:17128361	
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 Bachmanov, A.A., Beauchamp, G.K., Chatterjee, A., de Jong, P.J., Li, S., Li, X., Olmen, J.D., Reed, D.R., Ross, D., and Tordoff, M.G. 2001. Gene and sequence variation associated with sensing carbohydrate compounds and other sweeteners Patent: WO 0183749-A 4 08-NOV-2001; WANNER-LAMBERT COMPANY (US); The Monell Chemical Senses Center (US)		
JOURNAL			
FEATURES	Location/Qualifiers		
SOURCE	1..2559		
BASE COUNT	380 a 863 c 840 g 476 t		
ORIGIN	1..2559		
Query Match	99.9%; Score 2555.8; DB 6; Length 2559;		
Best Local Similarity	99.9%; Fred. No. 0;		
Matches 2557; Conservative	0; Mismatches 2; Indels 0; Gaps 0;		
QY	1	ATGCTGGGCGCTGCTGCTCTGCTGAGCCCTCTGAGGCTCTCTGACCTGAGGAG	60
DB	1	ATGCTGGGCGCTGCTGCTCTGCTGAGCCCTCTGAGGCTCTCTGACCTGAGGAG	60
QY	61	GGCCCATGTGTGCTGTACAGCACTTAGAGTGAAGGGGAACTACGTGCTGGGGGGCTG	120
DB	61	GGCCCATGTGTGCTGTACAGCACTTAGAGTGAAGGGGAACTACGTGCTGGGGGGCTG	120
QY	121	TTCGCCCTGGGAGAGCGAGAGAGGTGGCCCTCCGACGCGGACAGCGGCCAGACCT	180
DB	121	TTCGCCCTGGGAGAGCGAGAGAGGTGGCCCTCCGACGCGGACAGCGGCCAGACCT	180
QY	181	GTGTCCACACAGTTCTCTCAACAGGCTGTCTGTGGCACTGGCCATGAAATGGCCGTG	240
DB	181	GTGTCCACACAGTTCTCTCAACAGGCTGTCTGTGGCACTGGCCATGAAATGGCCGTG	240
QY	241	GAGGAGATCAACAAAGTCGATCTGCTGCCCGGGCTGCGCTTGCGCTAGACCTCTTT	300
DB	241	GAGGAGATCAACAAAGTCGATCTGCTGCCCGGGCTGCGCTTGCGCTAGACCTCTTT	300
QY	301	GATAGTGCTGGAGCGCTGTGGTGGCCATGAAGCCAGCGCTATGCTTCCTGGCCAAAGCA	360

Db 301 GATAGTGTCTGAGAGCTGTGGTGGCCATGAAGCCAGCTCATGTTCTTGCCGAAGGCA 360
QY 361 GGCAGCCCGGAGATAGCCGCTACTGTCACTACAGCACTACAGAGCCCGCTGTGCTGCT 420
Db 361 GGCAGCCCGGAGATAGCCGCTACTGTCACTACAGCACTACAGAGCCCGCTGTGCTGCT 420
QY 421 GTATATGGGCGCCACTGTGTCAAGAGCTGGCCATGTGTACCGGCAAGTTCTTACGCTTCTC 480
Db 421 GTATATGGGCGCCACTGTGTCAAGAGCTGGCCATGTGTACCGGCAAGTTCTTACGCTTCTC 480
QY 481 CTCATGCCCGAGGTACGTACAGTGTGTACATGAGAGCTGTGAGCGCCCGGAGAGACTTC 540
Db 481 CTCATGCCCGAGGTACGTACAGTGTGTACATGAGAGCTGTGAGCGCCCGGAGAGACTTC 540
QY 541 CCTCTCTTCTCGAGCGCTGGCCAGCGAGCGAGTGTACAGCGGCGCCGCGAGAGCTG 600
Db 541 CCTCTCTTCTCGAGCGCTGGCCAGCGAGCGAGTGTACAGCGGCGCCGCGAGAGCTG 600
QY 601 CTGCAAGAGTGTGGCTGGAATGGGTGGCCGCTGGGCGAGCGAGAGAGTACGCGCG 660
Db 601 CTGCAAGAGTGTGGCTGGAATGGGTGGCCGCTGGGCGAGCGAGAGAGTACGCGCG 660
QY 661 CAGGGGCTGAGCATCTTCTCGGCGCTGGCGACGCGGATCTGATCGCGACAGAG 720
Db 661 CAGGGGCTGAGCATCTTCTCGGCGCTGGCGACGCGGATCTGATCGCGACAGAG 720
QY 721 GGCCTGTGGCGCTGGCCGCTGCGATGATGCTGGCGCTGGGGAAGTGTAGAGAGCTCTG 780
Db 721 GGCCTGTGGCGCTGGCCGCTGCGATGATGCTGGCGCTGGGGAAGTGTAGAGAGCTCTG 780
QY 781 CACCAAGTGAACAGAGAGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db 781 CACCAAGTGAACAGAGAGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
QY 841 CAGCCCTCTTCACTACATGATCAAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db 841 CAGCCCTCTTCACTACATGATCAAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
QY 901 GAGGCGTGTGACCTGTGACCTGTGATGAGGAGGCTGGCGGATGGCCCAAGATGGGCAAG 960
Db 901 GAGGCGTGTGACCTGTGACCTGTGATGAGGAGGCTGGCGGATGGCCCAAGATGGGCAAG 960
QY 961 GTGCTTGTGCTCTCCAGAGAGGAGTGCACAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
Db 961 GTGCTTGTGCTCTCCAGAGAGGAGTGCACAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
QY 1021 CACCTGTGCGCTGTGCAAGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Db 1021 CACCTGTGCGCTGTGCAAGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
QY 1081 CTGAGAGAGAGAGT 1140
Db 1081 CTGAGAGAGAGAGT 1140
QY 1141 GTGAGAGAGAGAGT 1200
Db 1141 GTGAGAGAGAGAGT 1200
QY 1201 GCGCAGAGAGAGT 1260
Db 1201 GCGCAGAGAGAGT 1260
QY 1261 GTGAGAGAGAGT 1320
Db 1261 GTGAGAGAGAGT 1320
QY 1321 CCGCTGTGCTGT 1380
Db 1321 CCGCTGTGCTGT 1380
QY 1381 TGGCAGAGAGT 1440
Db 1381 TGGCAGAGAGT 1440

QY 1441 GAGCGCTGTGAATCCGCTGTGACACAGTGTGTGAACAACAGAACCCGCTGTGCTGCTGTG 1500
Db 1441 GAGCGCTGTGAATCCGCTGTGACACAGTGTGTGAACAACAGAACCCGCTGTGCTGCTGTG 1500
QY 1501 CGCAGTGTGAGAGAGAGT 1560
Db 1501 CGCAGTGTGAGAGAGAGT 1560
QY 1561 TGT 1620
Db 1561 TGT 1620
QY 1621 TGT 1680
Db 1621 TGT 1680
QY 1681 TTTCTGT 1740
Db 1681 TTTCTGT 1740
QY 1741 GGCCTGT 1800
Db 1741 GGCCTGT 1800
QY 1801 GGCCTGT 1860
Db 1801 GGCCTGT 1860
QY 1861 GTCTCTGT 1920
Db 1861 GTCTCTGT 1920
QY 1921 CACCTGT 1980
Db 1921 CACCTGT 1980
QY 1981 GAGT 2040
Db 1981 GAGT 2040
QY 2041 GGCCTGT 2100
Db 2041 GGCCTGT 2100
QY 2101 GT 2160
Db 2101 GT 2160
QY 2161 CACTGT 2220
Db 2161 CACTGT 2220
QY 2221 GGCCTGT 2280
Db 2221 GGCCTGT 2280
QY 2281 GGCCTGT 2340
Db 2281 GGCCTGT 2340
QY 2341 CTCTGT 2400
Db 2341 CTCTGT 2400
QY 2401 TGT 2460
Db 2401 TGT 2460
QY 2461 CAGAGGT 2520
Db 2461 CAGAGGT 2520

QY 2521 AATGACGGGAAACAGGAAATCAGGGGAAACATGATGA 2559
 Db 2521 AATGACGGGAAACAGGAAATCAGGGGAAACATGATGA 2559
 RESULT 4
 LOCUS AX239647 3489 bp DNA linear PAT 26-SEP-2001
 DEFINITION Sequence 13 from Patent WO0164882.
 ACCESSION AX239647
 VERSION AX239647.1 GI:15797320
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 AUTHORS Glucksmann, M.A., Galvin, K.M. and Silos-Santligo, I.
 TITLE 1 (bases 1 to 3489)
 JOURNAL 1983, 52881, 2398, 45449, 50289, and 52872, G protein-coupled
 receptors and uses therefor
 Patent: WO 0164882-A 13 07-SEP-2001;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source location/Qualifiers
 1..3489
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 52..2610
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC8308.1"
 /db_xref="GI:15797321"
 /translation="MLGPAYLIGLSLWALHPGTGAPLCLSQLRMKGDVYLGLPLPLG
 EAEAGLRSTRPSPVCTRFSSNGLLMALAMKMAVEINNKDLPGLGLDLPDT
 CSBPVAMKSLMFLAKAGSRDIAACNYOXPRLAVAGPSSSLAMTGRFSEF
 LMPQVSGASMEILSAREPSEFFRPVPSDVOYLAELQVFESEVMVVAALSDSEY
 GROLSTESLAARGICIAHEGLVPLPRADSDRGVGDVLOVNOVSSVOYLLEPS
 VHAHAFNYSISRLSPKWWASAMTSDIYMGIPGMOMTVIGETLORGLHEF
 PORYKTHLALATPAPCSALGEROGLEDVVGRCPCCTITLONVSAGLHHQTES
 VYAAVSVAAQALNTLQCNASGCPADQPVKPOLLEMYMLTFHVGLPLRFDSSGV
 RMYEDLKLVWQDSVPLRHVDVFNGLSLTERLIRMTDNKRPVRCRROCEGV
 RYKGFHSCEYDCVDEAGSYRONPDIACTFCGDSWSPERSRCRPSRIAMGE
 PAYLLLLLSLGLYALGLFVHHRDSPLVOASGGLACGIVGLGYCLSVLLE
 PGSPSPARCIAOPLSHPLTGLCTSLTFLDAALTPVSESLPMADRLSCINGPMAV
 LTVLLAMLVYALCTMTLVAPPEVYTDHMLPEALVHCRKTSWVSEGLAHATNATL
 AFLCFGLTFLVRSQPGRYNRAAGLITFAMLAFTTWSEVPLLANOVVILNPAYOMGL
 LILCVLIGLAFHPLPRCYLIMROPLMROPLNTPERFLGPGDAGOGDNGTNGOKEH"
 BASE COUNT 578 a 1189 c 1102 g 612 t 8 others
 ORIGIN
 Query Match 99.98; Score 2555.8; DB 6; Length 3489;
 Best Local Similarity 99.98; Pred. No. 0;
 Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 301 GATACGTCTGGAGACCTGTGTGTGGCCATGAGCCAGCCATGATCTGTGGCCAGGCA 360
 Db 352 GATACGTCTGGAGACCTGTGTGTGGCCATGAGCCAGCCATGATCTGTGGCCAGGCA 411
 QY 361 GGCAGCCGCGACATCGCGGCTACTGCACTACAGGAGTACAGCCCGGTGTGGCT 420
 Db 412 GGCAGCCGCGACATCGCGGCTACTGCACTACAGGAGTACAGCCCGGTGTGGCT 471
 QY 421 GTCATCGGGGCCCCACTGTGTAGAGCTGCGCATATGATACCCGCAAGTTCTTACGTTTTC 480
 Db 472 GTCATCGGGGCCCCACTGTGTAGAGCTGCGCATATGATACCCGCAAGTTCTTACGTTTTC 531
 QY 481 CTCATGCCCAAGGTACGCTACGCTACGCTGCTACATGAGTACGCTGAGCCCGGAGACCTTC 540
 Db 532 CTCATGCCCAAGGTACGCTACGCTGCTACATGAGTACGCTGAGCCCGGAGACCTTC 591
 QY 541 CCCCTCTTCTTCCGACCGGCCCCAGGAGCCGTGTGAGTGAACGGCCGCGGAGACTG 600
 Db 592 CCCCTCTTCTTCCGACCGGCCCCAGGAGCCGTGTGAGTGAACGGCCGCGGAGACTG 651
 QY 601 CTCGAGAGAGTTGCGCTGGAACCTGGGTGGCCGCTGGGCGAGCGAGAGTACGGCCGG 660
 Db 652 CTCGAGAGAGTTGCGCTGGAACCTGGGTGGCCGCTGGGCGAGCGAGAGTACGGCCGG 711
 QY 661 CAGGCGCTGAGCATCTTCTGGCCCTGGCCCGGCGAGCGCGCATCTGCGGCGAGAG 720
 Db 712 CAGGCGCTGAGCATCTTCTGGCCCTGGCCCGGCGAGCGCGCATCTGCGGCGAGAG 771
 QY 721 GGCCTGTGGCGGCTGCGCCCGGCGGATGACTGCGGGGTGGGGAAGGTGAGAGAGTCTGT 780
 Db 772 GGCCTGTGGCGGCTGCGCCCGGCGGATGACTGCGGGGTGGGGAAGGTGAGAGAGTCTGT 831
 QY 781 CACCAAGTGAACACAGACAGCGTGACAGGTGTGTGTCTGTTCGCTTCGTCACAGCCGCC 840
 Db 832 CACCAAGTGAACACAGACAGCGTGACAGGTGTGTGTCTGTTCGCTTCGTCACAGCCGCC 891
 QY 841 CAGCCCTCTTCAACTACATGATGAGAGAGAGCTCTCGGCCCAAGGTGTGGTGGCCAGC 900
 Db 892 CAGCCCTCTTCAACTACATGATGAGAGAGAGCTCTCGGCCCAAGGTGTGGTGGCCAGC 951
 QY 901 GAGGCTGTGGTGAACCTGTGACCTGTGATGAGGAGGTGGCCGATGGCCAGATGGGACG 960
 Db 952 GAGGCTGTGGTGAACCTGTGACCTGTGATGAGGAGGTGGCCGATGGCCAGATGGGACG 1011
 QY 961 GTGCTGTGGCTTCTTCAGAGAGGGGTGCCAGCTGCACAGATTCCCGCAGTACGTAAAGC 1020
 Db 1012 GTGCTGTGGCTTCTTCAGAGAGGGGTGCCAGCTGCACAGATTCCCGCAGTACGTAAAGC 1071
 QY 1021 CACCTGGCCCTTGGGACACGACCCGCGCTTGTGCTCTGCCCTGGGCGAGAGAGCAAGGT 1080
 Db 1072 CACCTGGCCCTTGGGACACGACCCGCGCTTGTGCTCTGCCCTGGGCGAGAGAGCAAGGT 1131
 QY 1081 CTGAGAGAGAGAGTGTGTGGGCGACGCTGCCCGAGTGTGACTGATACAGCTCAGCAAC 1140
 Db 1132 CTGAGAGAGAGAGTGTGTGGGCGACGCTGCCCGAGTGTGACTGATACAGCTCAGCAAC 1191
 QY 1141 GTGAGCGCAGGGCTAAATACACACAGAGCTTCTGTCTACGAGCTGTGTATAGCGTG 1200
 Db 1192 GTGAGCGCAGGGCTAAATACACACAGAGCTTCTGTCTACGAGCTGTGTATAGCGTG 1251
 QY 1201 GCCCAGGCGCTGCACAACTCTTCACTGCAAGGCTCAGAGCTGCCCGGCGAGAGCC 1260
 Db 1252 GCCCAGGCGCTGCACAACTCTTCACTGCAAGGCTCAGAGCTGCCCGGCGAGAGCC 1311
 QY 1261 GTGAGGCGCTGGAGCGCTGTGAGAAATGTACAACTGTGACCTTCCAGGTGGGGGGCTG 1320
 Db 1312 GTGAGGCGCTGGAGCGCTGTGAGAAATGTACAACTGTGACCTTCCAGGTGGGGGGCTG 1371
 QY 1321 CCGCTGCGGTTTCAGACAGCGGAAAGGTGAGATGAGTACGACGATGAAGCTGTGGGTG 1380
 Db 1372 CCGCTGCGGTTTCAGACAGCGGAAAGGTGAGATGAGTACGACGATGAAGCTGTGGGTG 1431
 QY 1381 TGGCAGGGGCTCATGTGCCAGAGCTTCACAGCTGGGGCAGTTTCAACGGCAGCTTCAGGACA 1440

|||||
Db 1432 TGGCAGGGCTCAGTGGCCAGGCTCCAGACAGTGGAGGTTCAACGGACCTCAGAGACA 1491
QY 1441 GAGGGCTGAAGATCCGGTGGGACAGTGTGACAAACGAAGGCCCTGTGCCGTCTCG 1500
Db 1492 GAGCGCTGAAGATCCGGTGGGACAGTGTGACAAACGAAGGCCCTGTGCCGTCTCG 1551
QY 1501 CGGCACTGGCAGAGGAGGAGGAGTGCAGGCGGATCAAGGGGTTCCATCTGCTCTACGAC 1560
Db 1552 CGGCACTGGCAGAGGAGGAGGAGTGCAGGCGGATCAAGGGGTTCCATCTGCTCTACGAC 1611
QY 1561 TGTGTGACTGCGAGGCGGGGACGTACCGGCAAAACCCAGAGACATCGCTGACACCTTT 1620
Db 1612 TGTGTGACTGCGAGGCGGGGACGTACCGGCAAAACCCAGAGACATCGCTGACACCTTT 1671
QY 1621 TGTGTGAGGATGATGTGTGTCCCGGAGCGAAGACAGCTGCTTCCGCGGAGGTCTCG 1680
Db 1672 TGTGTGAGGATGATGTGTGTCCCGGAGCGAAGACAGCTGCTTCCGCGGAGGTCTCG 1731
QY 1681 TTCTGGCATGGGGGAGGCGGCTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Db 1732 TTCTGGCATGGGGGAGGCGGCTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1791
QY 1741 GGCTGTGCTGCTGCTGCTTGGGGGCTTTCGTTTCAACATCGGGAGACCCACTGTTCAG 1800
Db 1792 GGCTGTGCTGCTGCTTGGGGGCTTTCGTTTCAACATCGGGAGACCCACTGTTCAG 1851
QY 1801 GGCTGGGGGGGGCCCTGGGCTGTGTGGCTGTGTGGCTGTGGCTGTGGCTGTGGCTGTG 1860
Db 1852 GGCTGGGGGGGGCCCTGGGCTGTGTGGCTGTGTGGCTGTGGCTGTGGCTGTGGCTGTG 1911
QY 1861 GTCTCTCTCTCTCTCTGCGGAGCGGACCGCTGCGGATGCTGCTGCTGCTGCTGCTGCTG 1920
Db 1912 GTCTCTCTCTCTCTCTGCGGAGCGGACCGCTGCGGATGCTGCTGCTGCTGCTGCTGCTG 1971
QY 1921 CACTCTCCGCTACGGGGTGTGCTGAGACACTTCTCTGAGGGGCGGCGGAGATCTTCGTG 1980
Db 1972 CACTCTCCGCTACGGGGTGTGCTGAGACACTTCTCTGAGGGGCGGCGGAGATCTTCGTG 2031
QY 1981 GAGTCAGAACTGCTGTGAGCTGGGAGACCGGGTGTGAGGGGCTGTGGGGGGGGGGCTG 2040
Db 2032 GAGTCAGAACTGCTGTGAGCTGGGAGACCGGGTGTGAGGGGCTGTGGGGGGGGGGCTG 2091
QY 2041 GGCTGTGCTGTGTGTGTGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2100
Db 2092 GGCTGTGCTGT 2151
QY 2101 GTGGCTTCCCGCGGAGGT 2160
Db 2152 GTGGCTTCCCGCGGAGGT 2211
QY 2161 CACTGCGGACAGCTGTGAGCTGGGTCAGCTTGTGAGCGGACCGGACCGGACCGGACCGG 2220
Db 2212 CACTGCGGACAGCTGTGAGCTGGGTCAGCTTGTGAGCGGACCGGACCGGACCGGACCGG 2271
QY 2221 GGCTTCTGT 2280
Db 2272 GGCTTCTGT 2331
QY 2281 GGCCGCGGCTACCTTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2340
Db 2332 GGCCGCGGCTACCTTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2391
QY 2341 CTCTCTGAGCAATGTGAGGT 2400
Db 2392 CTCTCTGAGCAATGTGAGGT 2451
QY 2401 TGTGTCTGAGCAATGT 2460
Db 2452 TGTGTCTGAGCAATGT 2511
QY 2461 CGAGGGCTCAACACCCCGAGTTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2520
|||||

Db 2512 CCAGGGCTCAACACCCCGAGTTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2571
QY 2521 AATGAGGGAGAACAGGAATTCAGGGGAAACAGAGTGA 2559
Db 2572 AATGAGGGAGAACAGGAATTCAGGGGAAACAGAGTGA 2610
RESULT 5
AX354026
LOCUS AX354026 2729 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 12 from Patent WO0203848.
ACCESSION AX354026
VERSION AX354026.1 GI:18618963
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 Naito A.T.
AUTHORS Method for opening the blood-brain barrier
TITLE Patent: WO 0203848-A 12 17-JAN-2002;
JOURNAL NAITO, Albert T. (US)
FEATURES
Location/Qualifiers
1..2729
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 7475194CB1"
BASE COUNT 413 a 934 c 879 g 503 t
ORIGIN
Query Match 98.3%; Score 2514.4; DB 6; Length 2729;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 2558; Conservative 0; Mismatches 1; Indels 33; Gaps 1;
QY 1 ATGCTGGGCTCTGCTGTCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 60
Db 1 ATGCTGGGCTCTGCTGTCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 60
QY 61 GCGCCATTTGGCTGTGACAGCAACTTAGATGAAGGGAGCACTGCTGTGGGGGGGGCTG 120
Db 61 GCGCCATTTGGCTGTGACAGCAACTTAGATGAAGGGAGCACTGCTGTGGGGGGGGCTG 120
QY 121 TTCCCTCTGGGCGAGGCGGAGGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 180
Db 121 TTCCCTCTGGGCGAGGCGGAGGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 180
QY 121 TTCCCTCTGGGCGAGGCGGAGGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 180
Db 121 TTCCCTCTGGGCGAGGCGGAGGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT 180
QY 181 GTGTGACCAAGGTTTCTCAAAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Db 181 GTGTGACCAAGGTTTCTCAAAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
QY 241 GAGGAATCAACAAGT 300
Db 241 GAGGAATCAACAAGT 300
QY 301 GATACGT 360
Db 301 GATACGT 360
QY 361 GGCAGCGGAGCATTCGCGCTGT 420
Db 361 GGCAGCGGAGCATTCGCGCTGT 420
QY 421 GTGATGGGGCGGACGT 480
Db 421 GTGATGGGGCGGACGT 480
QY 481 CTGATCCCCAGGTGCTAGCTGT 540
Db 481 CTGATCCCCAGGTGCTAGCTGT 540
QY 541 CCCTCTTCTTCCGACCGGTGCCAGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
|||||

Dh 541 CCGCTTCCTTCGCAACCGGCGGAGCGGCTGTGACGTGACGCGCGGAGGCTG 600
Qy 601 CTGCAAGAGTTCGGCTGGAACCTGGGTGGCCGCTGGGCGACGAGAGTACGGCCG 660
Dh 601 CTGCAAGAGTTCGGCTGGAACCTGGGTGGCCGCTGGGCGACGAGAGTACGGCCG 660
Qy 661 CAGGAGCTGAGCACTTCCTGGCCCTGGCCGCGGCGAGCGGCACTTCGATCGCGACAG 720
Dh 661 CAGGAGCTGAGCACTTCCTGGCCCTGGCCGCGGCGAGCGGCACTTCGATCGCGACAG 720
Qy 721 GGCCCTGGGCGGCTGGCCGCGGATGATCGGCGGCTGGGGAAGGTGACAGAGCTCTG 780
Dh 721 GGCCCTGGGCGGCTGGCCGCGGATGATCGGCGGCTGGGGAAGGTGACAGAGCTCTG 780
Qy 781 CACCAAGTGAACAGAGAGCGGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Dh 781 CACCAAGTGAACAGAGAGCGGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Qy 841 CACGCGCTTCCTCACTACATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Dh 841 CACGCGCTTCCTCACTACATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Qy 901 GAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 960
Dh 901 GAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 960
Qy 961 GTGCTTGGCTTCCTCCAGAGGGGTGCCAGCTGCAAGAGTTCCTCCAGTACGTGAAGAG 1020
Dh 961 GTGCTTGGCTTCCTCCAGAGGGGTGCCAGCTGCAAGAGTTCCTCCAGTACGTGAAGAG 1020
Qy 1021 CACCTGGGCTTCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
Dh 1021 CACCTGGGCTTCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1080
Qy 1081 CTGAGAGAGAGAGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
Dh 1081 CTGAGAGAGAGAGTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
Qy 1141 GTGAGAGCGAGGGCTTAATCACCACAGAGCTTCCTGTCTAGCGAGCTGTATAGCGTG 1200
Dh 1141 GTGAGAGCGAGGGCTTAATCACCACAGAGCTTCCTGTCTAGCGAGCTGTATAGCGTG 1200
Qy 1201 GCGGAGGCGGCTGCAACACTTCCTAGTGAAGCGCTGAGGCTGCGGCGGCGGCGGCGG 1260
Dh 1201 GCGGAGGCGGCTGCAACACTTCCTAGTGAAGCGCTGAGGCTGCGGCGGCGGCGGCGG 1260
Qy 1261 GTGAAGCGGCTGAGCTTCCTGAGAGAACTGTACAACTGACCTTCACAGTGGGCGG 1320
Dh 1261 GTGAAGCGGCTGAGCTTCCTGAGAGAACTGTACAACTGACCTTCACAGTGGGCGG 1320
Qy 1321 CCGCTGGGCTTCGACAGAGCGGGAACGTGAGATGAGTGAAGCTGTGGGTG 1380
Dh 1321 CCGCTGGGCTTCGACAGAGCGGGAACGTGAGATGAGTGAAGCTGTGGGTG 1380
Qy 1381 TGGAGAGGCTTCAGTGGCCAGGCTTCAGAGAGCTGGGAGGTTCAACGGGAGCTCAG 1440
Dh 1381 TGGAGAGGCTTCAGTGGCCAGGCTTCAGAGAGCTGGGAGGTTCAACGGGAGCTCAG 1440
Qy 1441 GAGGAGCTTCAGATTCCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTG 1476
Dh 1441 GAGGAGCTTCAGATTCCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTG 1476
Qy 1477 -----CAGAGAGCGGCTTCCTGGGTGCTGGGCGGAGTGCAGAGAGGCGGAGTGGC 1527
Dh 1501 GCGTGTGGCGAGAGCGGCTTCCTGGGTGCTGGGCGGAGTGCAGAGAGGCGGAGTGGC 1560
Qy 1528 CCGGCTCAAGGGGCTTCCTGCTGCTAGAGAGTGTGGAGTGTGGAGGCGGCGAGCTAC 1587
Dh 1561 CCGGCTCAAGGGGCTTCCTGCTGCTAGAGAGTGTGGAGTGTGGAGGCGGCGAGCTAC 1620
Qy 1588 CCGGCAAAACCGAGAGAGTGTGGAGCTTCCTGAGAGTGTGGAGTGTGGAGTGTGGAG 1647
Dh 1621 CCGGCAAAACCGAGAGAGTGTGGAGCTTCCTGAGAGTGTGGAGTGTGGAGTGTGGAG 1680

Qy 1648 CGAAGCACACGCTGCTTCCTGGCGGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1707
Dh 1681 CGAAGCACACGCTGCTTCCTGGCGGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Qy 1708 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1767
Dh 1741 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
Qy 1768 TTGCTTCACATCGGAGAGAGCCACTGATGAGAGCTTCGAGAGGAGGAGGAGGAGGAG 1827
Dh 1801 TTGCTTCACATCGGAGAGAGCCACTGATGAGAGCTTCGAGAGGAGGAGGAGGAGGAG 1880
Qy 1828 GGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1887
Dh 1861 GGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
Qy 1888 CCGGCGGAGTGCCTGGGCGGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1947
Dh 1921 CCGGCGGAGTGCCTGGGCGGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
Qy 1948 ACATCTTCCTGAGAGGCGGCGGAGATCTGCTGAGTGAAGTGAAGTGAAGTGAAGTGA 2007
Dh 1981 ACATCTTCCTGAGAGGCGGCGGAGATCTGCTGAGTGAAGTGAAGTGAAGTGAAGTGA 2040
Qy 2008 GACCGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2067
Dh 2041 GACCGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Qy 2068 CTGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2127
Dh 2101 CTGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2160
Qy 2128 GACTGTGACATGCTGCCAGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2187
Dh 2161 GACTGTGACATGCTGCCAGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Qy 2188 TTGCGGCTGAGCGAGCGGCGGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2247
Dh 2221 TTGCGGCTGAGCGAGCGGCGGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
Qy 2248 CTGCTGAGAGGAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2307
Dh 2281 CTGCTGAGAGGAGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
Qy 2308 GCTACTTTCATCACTGAGGCTTCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2367
Dh 2341 GCTACTTTCATCACTGAGGCTTCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
Qy 2368 AGGCGCGGCTGAGATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2427
Dh 2401 AGGCGCGGCTGAGATGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
Qy 2428 CACCTGCGAGAGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2487
Dh 2461 CACCTGCGAGAGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
Qy 2488 CTGGAGAGGAGGCGGCTGAGGAGTGCCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 2547
Dh 2521 CTGGAGAGGAGGCGGCTGAGGAGTGCCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAG 2580
Qy 2548 AAACATGAGTGA 2559
Dh 2581 AAACATGAGTGA 2592

RESULT 6
AX318243
LOCUS AX318243 2739 bp DNA linear PAT 14-DEC-2001
DEFINITION Sequence 24 from Patent WO0190359.
ACCESSION AX318243
VERSION AX318243.1 GI:17900903
KEYWORDS

OY	1744	CTGTGTCGCGGCTTTTGGGCGCTTCCATCAACATCCGGGACAGGCCACTGGTTCAAGCC	1803
Db	1777	CTTGTGTGGCTCTTTGGGCGTGTTCATCAACATCCGGGACAGGCCACTGGTTCAAGCC	1836
OY	1804	TCGGGGGGGGCCCTTGCCCTGTGCTTTTGGCCCTGTGTCTCTGTGGGCTGTGCTCAAGCTC	1863
Db	1837	TCGGGGGGGGGGCCCTTGCCCTGTGCTTTTGGCCCTGTGTCTCTGTGGGCTGTGCTCAAGCTC	1896
OY	1864	CTTCTGTTCCTTGCCAGCCCAAGCCCTGCGATGCTGTGGCCAGACAGCCCTTGTCCAC	1923
Db	1897	CTTCTGTTCCTTGCCAGCCCAAGCCCTGCGATGCTGTGGCCAGACAGCCCTTGTGTCCAC	1956
OY	1924	CTCCCGCTACAGGGGTGGCTGTAGCACACTTCCTCGAAGGGGGCGAGATCTTGTGGAG	1983
Db	1957	CTTCCGGCTACAGGGGTGGCTGTAGCACACTTCCTGTGACAGCGGGCCGAGATCTTGTGGAG	2016
OY	1984	TCAGAACTGCTCTGTAGCTGGGACAGCCGACTGTGTGCTGCTCGGGGGGCGCTTGAGCC	2043
Db	2017	TCAGAACTGCTCTGTAGCTGGGACAGCCGACTGTGTGCTGCTCGGGGGGCGCTTGAGCC	2076
OY	2044	TGGCTGTGTGTCTCTGTGCCATGCTGTGTGAGAGTTCGCACTGTGCACCTGGTACTGTGTG	2103
Db	2077	TGGCTGTGTGTCTCTGTGCCATGTGTGTGAGAGTTCGCACTGTGCACCTGTGTACTGTGTG	2136
OY	2104	GCTTTCGCCCGCGAGGTGGTGAC--GGACTGGGACATGTGCGCCACGAGAGGCGCTGTGTGA	2162
Db	2137	GCTTTCGCCCGCGAGGTGGTGAC--GGACTGGGACATGTGCGCGCCACGAGAGGCGCTGTGTGA	2196
OY	2163	CTGCGGCACACGCTCTGTGGGTAGGTTGCGCCCTAGCGCAAGCCACCAATGCCACGCTGGC	2222
Db	2197	CTGCGGCACACGCTCTGTGGGTAGGTTGCGCGGCTTAGCGCACAGCGCACCC--ATGCGACGCTGGC	2255
OY	2223	CTTTCCTGTGCTTC--CTGGGCACTTTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2281
Db	2256	CTTTCCTGTGCTTCCTGTGGCACTTTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2315
OY	2282	CC---CGTGGCCCTACCTTTGGCATGTGCGGCTACTTTCATCTAC--CTGGGCTTC--CTTGT	2336
Db	2316	CCAGGTGGGCTTACACATTTTGGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2375
OY	2337	GCCCTCTCTGGC--CAATGTGCAGGTGGT--CCTACAGGCCCGCTGTGCAGATGGGCGCCT	2393
Db	2376	GCCCTCTCTCTGGCAGATGTGCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2435
OY	2394	CTGTCTGT	2453
Db	2436	CTGTCTGT	2495
OY	2454	GGGCGACCCAGGGCTCACACACCCCGAGTTCTTCTGTGGAGGGGGCGCTGTGGGATGCC--C	2512
Db	2496	GGGCGACCCAGGGCTCACACACCCCGAGTTCTTCTGTGGAGGGGGCGCTGTGGGATGCCAC	2555
OY	2513	AAGGCGAATGTAGGGGAACAGAGAAATCAGGGGAAACATAGTGA	2559
Db	2556	AAGGCGAATGTAGGGGAACAGAGAAATCAGGGGAAACATAGTGGTGA	2602

	RESULT 7	
LOCUS	AB065647	
DEFINITION	AB065647 Homo sapiens gene for seven transmembrane helix receptor, complete cds, isolate:CBRC7TM_210.	4256 bp DNA linear PRI 23-JUL-2002
ACCESSION	AB065647	
VERSION	AB065647.1 GI:21928568	
KEYWORDS	.	
SOURCE	Homo sapiens (isolate:CBRC7TM_210) DNA.	
ORGANISM	Homo sapiens	
	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 Suwa,M., Sato,T., Okouchi,I., Arita,M., Futami,K., Matsumoto,S., Tsutsumi,S., Aburatani,H., Asai,K. and Akiyama,Y. Genome-wide discovery and analysis of human seven transmembrane	
AUTHORS		
TITLE		

JOURNAL helix receptor genes
 REFERENCE Unpublished
 2 (bases 1 to 4256)
 AUTHORS Suwa, M.
 TITLE Direct Submission
 JOURNAL Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST), 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan (E-mail: m-suwa@aist.go.jp, URL: http://www.cbrc.jp/, Tel: 81-3-3599-8080, Fax: 81-3-3599-8081)
 COMMENT This sequence is a seven transmembrane helix receptor candidate predicted from the whole human genome sequences using our automated system that contains programs of gene finding (GeneDecoder), sequence search, motif-domain assignment and transmembrane helix prediction.
 [Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST)] and [Genome Science Division, Research Center for Advanced Science and Technology (KCAST), University of Tokyo].
 FEATURES
 source Location/Qualifiers
 1..4256
 /organism="Homo sapiens"
 /isolate="CBRC7TM.210"
 /db_xref="taxon:9606"
 /chromosome="1"
 join(201..391,493..793,879..1661,1776..1979,2114..223361..3315,3713..3890,3979..4056)
 /codon_start=1
 /evidence="not_experimental"
 /product="seven transmembrane helix receptor"
 /protein_id="BAC05873.1"
 /db_xref="GI:21928569"
 /translation="MLGPVVLGLMALHPGIGAPLCSQOLRMKGDVYLIGFPLGSAEAGKSRPTSPVCTRRSSNGMLAMKMAVEINNKSLPLRGXDLPFTCEAEVVMKPSRTMFLAKGRSDIAACNTAYQYORVLAIVGSHSETAMTKRFEFLMPVYSGASHELLSARETPSFRRYRPSDRVQLTAALIELDEGRNVMVAALSDDYGRGLSTFSAALAAARGICIAHEGLVPLPRADDSRLGVODLVHVNOSVQVALLFASVHAALFENYSISSRLSPKPAVSEAMLTSDLVGLGMAOMGVIAFLDRQADLHFEPOVYRTHALATDPAFCSSALGEROGLIEDVVGRCPCDITLQNTASALINHQTSPYAAVYSVALAHNTLOCNAGSCPAODPKVMOLENNYILTFHGGILPRFDSGANVDMEDIKLMWOGSVPLRHVDGRNGSLRFRRLKIRHNTSNOKPVSRCMOCEQOVRKRVGHSSCTCYDCERGSRRQRPDIACFTGCDMSPSRSTRCRPSRPLAMEBPVAILLLLSLALGLVLAALGLVHHRDSLVAGSPLACGLVCLGLVSLVLPFGQSPAPACLAQPLSHLPLTGCLSTLEFLAAELFVSESLPLSWADRLSGLRQPMWLVVLIAMLEVVALCTWLVAFPEPVYDMMHLPTEALVHCRTBSWSFGLAHATNALLAFCLFETFLRSPGCGYNRARGLTAMLVAFITWVSFVPLAAVOVLPAYOMGLALLCVGLITAAFLHPRCYLIMRQPLINPEFLIGGPDGAGONDNGTNGNGKHEAPLILHPAAHSHSGRGCGTQAAVAGCARRGRGQPLPLPAAQWDSIMVTSIRDRADQDLTPARQAHHTGSPVRRPGSAPRA"

QY 188 -----CAGGTTCTGCT 199
DB 441 GTGCTCTGAGCTGGGGCCGAGAGTGGCATCTGCGGTTCTGTGTGCCCCAGAGTTCTCT 500
QY 200 CAAAGGCGCTGCTTGGGCACTGGCCATGAAATGGCCGTGAGAGATCAACAACAGT 259
DB 501 CAAAGGCGCTGCTTGGGCACTGGCCATGAAATGGCCGTGAGAGATCAACAACAGT 560
QY 260 CGGATCTGCTGGCGGGGCTGGGCGCTAGAGCTCTTTGATACGTCTGTGAGGCTG 319
DB 561 CGGATCTGCTGGCGGGGCTGGGCGCTAGAGCTCTTTGATACGTCTGTGAGGCTG 620
QY 330 TGTGTGCAATGAAGCCAGACCTCATGTTCTGTGGCCAGGAGGAGGAGCCGCAATCGCG 379
DB 621 TGTGTGCAATGAAGCCAGACCTCATGTTCTGTGGCCAGGAGGAGGAGCCGCAATCGCG 680
QY 380 CCTACTGCAATCAACGAGATACAGAGCCCGTGTGCTGCTGCTCATCGGGCCCACTGCT 439
DB 681 CCTACTGCAATCAACGAGATACAGAGCCCGTGTGCTGCTGCTCATCGGGCCCACTGCT 740
QY 440 CAGAGCTGGCCATGCTGACCGGCGAGTTCTTCAAGCTTCTGCTCATGCGCC----- 489
DB 741 CAGAGCTGGCCATGCTGACCGGCGAGTTCTTCAAGCTTCTGCTCATGCGCCAGTGGGCG 800
QY 490 ----- 489
DB 801 CCCCCACATCACCCACCACCAGCCCTGCCCCGTGGAGGCCCTGTGTGACGAGAT 860
QY 490 -----CAGTCACTAGAGTGTGCTAGCATGAGAGTCTGAGCGCCGCGAG 534
DB 861 GCGCTTGGCCCTTTCAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 920
QY 535 ACCTTCCCTCTCTTCTTCCGACGCTGCGCCAGCGACCGGTCAGCTGACGCGCGCGCG 594
DB 921 ACCTTCCCTCTCTTCTTCCGACGCTGCGCCAGCGACCGGTCAGCTGACGCGCGCGCG 980
QY 595 GAGCTGTGACGAGAGTTCGGCTGGAATGTGGTGGCCGCTTGGGAGCGAGCAGCAGATAC 654
DB 981 GAGCTGTGACGAGAGTTCGGCTGGAATGTGGTGGCCGCTTGGGAGCGAGCAGCAGATAC 1040
QY 655 GCGCGGCGAGGCTGAGATCTTCTCGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 714
DB 1041 GCGCGGCGAGGCTGAGATCTTCTCGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1100
QY 715 CAGGAGGCGCTGTGCTGCG 774
DB 1101 CAGGAGGCGCTGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1160
QY 775 GTCTGTGACACGAGTACACGAGCAGCGTGCAGTGTGTGCTGTGCTGCTGCTGCTGAC 834
DB 1161 GTCTGTGACACGAGTACACGAGCAGCGTGCAGTGTGTGCTGTGCTGCTGCTGCTGAC 1220
QY 835 GCGCGGCGAGGCTTCTTCACTACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 894
DB 1221 GCGCGGCGAGGCTTCTTCACTACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1280
QY 895 GCGAGCAGGCGCTGTGCTGACCTGTGACCTGTGATGGGCGTGCAGGATGGCCAGATG 954
DB 1281 GCGAGCAGGCGCTGTGCTGACCTGTGATGGGCGTGCAGGATGGCCAGGATG 1340
QY 955 GCGAGCAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014
DB 1341 GCGAGCAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1400
QY 1015 AAGAGCAGCAGGCGCTTGGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1074
DB 1401 AAGAGCAGCAGGCGCTTGGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1460
QY 1075 CAGGCTGTGAGAGAGAGTGTGGGCGAGCGCTGCCGAGTGTGATGATGATGATGATGATG 1134
DB 1461 CAGGCTGTGAGAGAGAGTGTGGGCGAGCGCTGCCGAGTGTGATGATGATGATGATGATG 1520
QY 1135 CAGAACTGTAGAGCGAGGCTTAATTCACACAGAGAGCTTCTGTGTGCTGACGAGCTGTAT 1194

DB 1521 CAGAACTGTAGAGCGAGGCTTAATTCACACAGAGAGCTTCTGTGTGCTGACGAGCTGTAT 1580
QY 1195 AGCGTGGCCAGAGCCCTGCACAACTCTTCAAGTCAAGAGCCCTGAGCTGCCCGCGAG 1254
DB 1581 AGCGTGGCCAGAGCCCTGCACAACTCTTCAAGTCAAGAGCCCTGAGCTGCCCGCGAG 1640
QY 1255 GACCCCGTGAAGCCCTG----- 1271
DB 1641 GACCCCGTGAAGCCCTGGAAGTGAAGCCCGGAGATGGGGGTGTGTCTCTGTGATGT 1700
QY 1272 ----- 1271
DB 1701 GCCCAGGCGACAGGCGACGCGCACAGCGCTGAGCTGAGAGTGGCTGGCGGCTCAGCCCC 1760
QY 1272 -----GACGCTCTGTGAGAACTGTCAACTGTGACCTTCCAGCTGGCGGCTG 1320
DB 1761 GTCCCGCCCGCAGCTCTGTGAGAACTGTCAACTGTGACCTTCCAGCTGGCGGCTG 1820
QY 1321 CCGCTGGGCTTGCAGAGAGCGGAAAGTGGACATGAGATGAGCTGAGCTGAGCTGTGGGTG 1380
DB 1821 CCGCTGGGCTTGCAGAGAGCGGAAAGTGGACATGAGATGAGCTGAGCTGAGCTGTGGGTG 1880
QY 1381 TGGCAGGCGCTAGTCCAGAGCTCCAGCAGCTGTGGCAGGTTCAAGCGAGCTCAGAGACA 1440
DB 1881 TGGCAGGCGCTAGTCCAGAGCTCCAGCAGCTGTGGCAGGTTCAAGCGAGCTCAGAGACA 1940
QY 1441 GAGCGCTGTGAAGATCCGCTGGCAGACGTTCTGACAAAC----- 1476
DB 1941 GAGCGCTGTGAAGATCCGCTGGCAGACGTTCTGACAAACAGTGTGAGTGTGGGTGTGC 2000
QY 1477 ----- 1476
DB 2001 CAGGCTGTGCTGTGTAGCCCCCGCGCGAGGCGCAGCTGGGGGTGGGGCCCTGTCCAG 2060
QY 1477 -----CAGAAAGCCG 1486
DB 2061 TCTCCGTGGGATGCCAGCGAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2120
QY 1487 TGTCCCGGCTGCGGCGAGTGTCCAGAGAGGCGCAGGTCGCGCGGCTCAAGGGCTTCCAT 1546
DB 2121 TGTCCCGGCTGCGGCGAGTGTCCAGAGAGGCGCAGGTCGCGCGGCTCAAGGGCTTCCAT 2180
QY 1547 CCTGCTGTAGACTGTGTGACTGCGAGGCGGCGAGTACCGCGCAAAAC----- 1597
DB 2181 CCTGCTGTAGACTGTGTGACTGCGAGGCGGCGAGTACCGCGCAAAACAGGTGAGC 2240
QY 1598 ----- 1597
DB 2241 CGCCTTCCCGCAGGCGGGGGGTGGGAACGACAGGAGGAGGCTGCCAAGTCTGACT 2300
QY 1598 -----CAG 1600
DB 2301 CTGAGACAGAGCCACAGAGGTACAAAGAACACACCCAGCGCCCTTCTCTCTCAGAG 2360
QY 1601 ACAGATGCGCTGACACTTTTGTGCGCAGAGATGATGTGTCCCGGAGGAGCAGACAGCT 1660
DB 2361 ACAGATGCGCTGACACTTTTGTGCGCAGAGATGATGTGTCCCGGAGGAGCAGACAGCT 2420
QY 1661 GCTTCCCGCAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1720
DB 2421 GCTTCCCGCAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2480
QY 1721 TGTGCTGTAGGCTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1780
DB 2481 TGTGCTGTAGGCTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2540
QY 1781 GGGACAGCCCACTGTGTAGGCTGTGCGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1840
DB 2541 GGGACAGCCCACTGTGTAGGCTGTGCGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2600
QY 1841 TGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1900

Db 31779 CCTACTGCAACTACAGAGATACAGACCCCGTGTGCTGCTGATCAGGCCCCACTCGT 31838
QY 440 CAGACTGCGCATATGTCACCGGCAATTTCTTACTTCTTCTCTATGCC----- 489
Db 31839 CAGACTGCGCATATGTCACCGGCAATTTCTTACTTCTTCTCTATGCCAGGTGGCG 31898
QY 490 ----- 489
Db 31899 CCCCCACCATACACCCACCCACCCGCTGGCCGTGGAGCCCTGTGTACAGAGAT 31958
QY 490 -----CAGTACGATACGATGCTATGATGAGCTCTGAGCCCGGAG 534
Db 31959 GCGCTTGGCCCTTACAGTACGATGCTATGAGCTCTGAGCTGAGCCCGGAG 32018
QY 535 ACCTTCCCTCTCTTCTTCCGACCGCTGGCCACGACCGGTGTACCTACAGCCCGCGG 594
Db 32019 ACCTTCCCTCTCTTCTTCCGACCGCTGGCCACGACCGGTGTACCTACAGCCCGCGG 32078
QY 595 GAGCTGCTGACAGAGTGGCTGGAATGGGTGGCCGCTGGGAGCGAGCAGAGATAC 654
Db 32079 GAGCTGCTGACAGAGTGGCTGGAATGGGTGGCCGCTGGGAGCGAGCAGAGATAC 32138
QY 655 GCGCGGACAGGCTGAGCATCTTCTGCGCCCTGGCCGCGGACGCGCATCTGATCGCG 714
Db 32139 GCGCGGACAGGCTGAGCATCTTCTGCGCCCTGGCCGCGGACGCGCATCTGATCGCG 32198
QY 715 CACGAGGCGCTGGTGGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCC 774
Db 32199 CACGAGGCGCTGGTGGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCC 32258
QY 775 GTCTGACACAGTGAACACAGAGCAGTGCAGTGTGCTGCTGTGCTGCTGCTGCTGAC 834
Db 32259 GTCTGACACAGTGAACACAGAGCAGTGCAGTGTGCTGCTGTGCTGCTGCTGCTGAC 32318
QY 835 GCGCGCCACAGCCCTTCTTACATACAGCATCAGCAGAGGCTTCCGCCAAGGTGTGGTG 894
Db 32319 GCGCGCCACAGCCCTTCTTACATACAGCATCAGCAGAGGCTTCCGCCAAGGTGTGGTG 32378
QY 895 GCCACGAGAGGCTTGGCTGCTGACCTGATGAGGCTGAGGCGCTGCGGCGATGGCCAGATG 954
Db 32379 GCCACGAGAGGCTTGGCTGCTGACCTGATGAGGCTGAGGCGCTGCGGCGATGGCCAGATG 32438
QY 955 GGCACGAGTGTGGCTTCTCTCCAGAGGGGTGCCAGCTGCAGAGATTCGCCCAAGTACGTG 1014
Db 32439 GGCACGAGTGTGGCTTCTCTCCAGAGGGGTGCCAGCTGCAGAGATTCGCCCAAGTACGTG 32498
QY 1015 AAGACGACACTTGGCCCTGGCCACCGAGCCGCTTCTGCTGCTGCTGGCGAGAGGAG 1074
Db 32499 AAGACGACACTTGGCCCTGGCCACCGAGCCGCTTCTGCTGCTGCTGGCGAGAGGAG 32558
QY 1075 CAGGGTCTGAGAGAGAGTGGTGGGCGACGCTGCCGACATGATGATCAACGCTG 1134
Db 32559 CAGGGTCTGAGAGAGAGTGGTGGGCGACGCTGCCGACATGATGATCAACGCTG 32618
QY 1135 CAGAACTGAGGCGAGGGCTTAATACACACAGAGCTTCTGCTTACGACAGCTGTAT 1194
Db 32619 CAGAACTGAGGCGAGGGCTTAATACACACAGAGCTTCTGCTTACGACAGCTGTAT 32678
QY 1195 AGCGTGGCCAGGCGCTTGCACAACTCTTCACTGACAGCGCTGAGGCTGCGCGGAG 1254
Db 32679 AGCGTGGCCAGGCGCTTGCACAACTCTTCACTGACAGCGCTGAGGCTGCGCGGAG 32738
QY 1255 GACCCCGTGAAGCCCTG----- 1271
Db 32739 GACCCCGTGAAGCCCTTGGAGGTGAGCCCGGAGATGGGGGTGTGCTGTCTGTGATGT 32798
QY 1272 ----- 1271
Db 32799 GCCACGAGCCACAGGACGCGCCAGTGTGAGTGGGTGGCGGCTCAGCCCC 32858
QY 1272 -----GCAGCTCTGTGAGAAATGTACAACTTACACTTGCAGTGGCGGCTG 1320
Db 32859 GTCCCGCGCGAGCTCTGTGAGAAATGTACAACTTACACTTGCAGTGGCGGCTG 32918

QY 1321 CCGGTGGGTTTCAGAGAGGAGGAAACGTGACATGGATGACACTGAGCTGTGGTG 1380
Db 32919 CCGGTGGGTTTCAGAGAGGAGGAAACGTGACATGGATGACACTGAGCTGTGGTG 32978
QY 1381 TGGCAGGAGCTCAGTGGCCAGAGCTCCACAGCTGGGCGAGTTCAAGCGACCTCAGACA 1440
Db 32979 TGGCAGGAGCTCAGTGGCCAGAGCTCCACAGCTGGGCGAGTTCAAGCGACCTCAGACA 33038
QY 1441 GAGCGCTGAGATCCGCTGGCAGACGTTCTGACAA----- 1476
Db 33039 GAGCGCTGAGATCCGCTGGCAGACGTTCTGACAAACAGTGTAGAGGTGGGTGTGC 33098
QY 1477 ----- 1476
Db 33099 CAGCGGTGCCGTGTGATGCCCGCGGACAGGGCGACAGCTGGGGGTGGGGCCGTTCCAG 33158
QY 1477 -----CAGAAAGCCG 1486
Db 33159 TCTCCGTTGGGATGCGCAGCGAGAGAGCCAGACCCAGGCTGTGTGGCAGAAAGCCG 33218
QY 1487 TGTCCCGGTGTCTCGGCGAGTCCACAGAGAGGCTCCAGGTGCGCGGTCAAGGGGTTCCACT 1546
Db 33219 TGTCCCGGTGTCTCGGCGAGTCCACAGAGAGGCTCCAGGTGCGCGGTCAAGGGGTTCCACT 33278
QY 1547 CCTGCTGATGACATGTGTGAGCTGCGAGGCGGCGAGCTACCGGGGAAAC----- 1597
Db 33279 CCTGCTGATGACATGTGTGAGCTGCGAGGCGGCGAGCTACCGGGGAAACCCAGGTGAGC 33338
QY 1598 ----- 1597
Db 33339 CCGCTTCCCGGAGGCGGGGGTGGGAAAGCAGCAGGAGGATCGTCCGAATCGTACT 33398
QY 1598 -----CAG 1600
Db 33399 CTGAGACAGAGCCCAAGAGGTACAAAGAAACACCAGCGCTTCTCTCTACAG 33458
QY 1601 ACAGATGCTGCTGACCTTTTGTGGCAGAGATGAGTGGTCCCGGAGCAGACAGCT 1660
Db 33459 ACAGATGCTGCTGACCTTTTGTGGCAGAGATGAGTGGTCCCGGAGCAGACAGCT 33518
QY 1661 GCTTCCGCGCAGAGTCTGCTGCTGCTGAGAGGGGCGAGCGGCTGTCTGTCTGCTCC 1720
Db 33519 GCTTCCGCGCAGAGTCTGCTGCTGCTGAGAGGGGCGAGCGGCTGTCTGTCTGCTCC 33578
QY 1721 TCTGCTGAGCTTGGCGCTTGTGCTGTGCTGTGCTTGTGGGCTGTCTGCTACATC 1780
Db 33579 TCTGCTGAGCTTGGCGCTTGTGCTGTGCTGTGCTTGTGGGCTGTCTGCTACATC 33638
QY 1781 GGGACAGCCACTGGTTCAGGCTCGGGGGGGCCCTGGGCTGCTTGGGCTGGTGTGC 1840
Db 33639 GGGACAGCCACTGGTTCAGGCTCGGGGGGGCCCTGGGCTGCTTGGGCTGGTGTGC 33698
QY 1841 TGGGCTGTGTCTGCTCAGAGCTCTCTGCTTCTGCTGCGCAGCCAGCCCTGCGCATGCG 1900
Db 33699 TGGGCTGTGTCTGCTCAGAGCTCTCTGCTTCTGCTGCGCAGCCAGCCCTGCGCATGCG 33758
QY 1901 TGGCCCAAGAGCCCTTGTCTCCACCTCCGCTTACGAGGCTGCTGAGACACTTCTCTGC 1960
Db 33759 TGGCCCAAGAGCCCTTGTCTCCACCTCCGCTTACGAGGCTGCTGAGACACTTCTCTGC 33818
QY 1961 AGCGGCGAGATCTTCTGAGAGTCAAGAACTCTCTGAGCTGGGAGAGCCGCTGAGTG 2020
Db 33819 AGCGGCGAGATCTTCTGAGAGTCAAGAACTCTCTGAGCTGGGAGAGCCGCTGAGTG 33878
QY 2021 GCTGCTGTGGGGGCGCTGGGCTGTGGTGTGTGTCTGTGCGCATGTGTGTGAGAGTGC 2080
Db 33879 GCTGCTGTGGGGGCGCTGGGCTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 33938
QY 2081 CACTGTGACCTGTGTACCTGTGTGGCTTTCGCCGAGGAGTGTGTGACGACTGGCAGATGC 2140
Db 33939 CACTGTGACCTGTGTACCTGTGTGGCTTTCGCCGAGGAGTGTGTGACGACTGGCAGATGC 33998

```

QY 2141 TGCCACGAGGCGCTGTGCTACGCGCACACGCTCCTGGGTAGCTTGGCCCTAGCGC 2200
Db 33999 TGCCACGAGGCGCTGTGCTACGCGCACACGCTCCTGGGTAGCTTGGCCCTAGCGC 34058
QY 2201 ACGCCACCAATGCGCAGCTGTGCTTCTGTCTCTCTGGGACACTTCTGTGGGAGCC 2260
Db 34059 ACGCCACCAATGCGCAGCTGTGCTTCTGTCTCTGGGACACTTCTGTGGGAGCC 34118
QY 2261 AGCCGCGCTGTACACACGCGCCGCGCCCTACCTTGGCATCTGTGGCTACTTCAATCA 2320
Db 34119 AGCCGCGCTGTACACACGCGCCGCGCCCTACCTTGGCATCTGTGGCTACTTCAATCA 34178
QY 2321 CTTGGGCTCTCTTGTGTGCTCTCTCTGCGCAATGTGACAGTGTCTTCAAGCCCGCTGC 2380
Db 34179 CTTGGGCTCTCTTGTGTGCTCTCTCTGCGCAATGTGACAGTGTCTTCAAGCCCGCTGC 34238
QY 2381 AGATGGGCGCGCTGTGCTGTGCTGTGCGCATCTGGGCGCTTCCACCTGCGCAGGT 2440
Db 34239 AGATGGGCGCGCTGTGCTGTGCTGTGCGCATCTGGGCGCTTCCACCTGCGCAGGT 34298
QY 2441 GTTACTGTCTCATGCGGCGAGCGCTCAACACCCCGAGTTCTTCTGGGAGGCGCC 2500
Db 34299 GTTACTGTCTCATGCGGCGAGCGCTCAACACCCCGAGTTCTTCTGGGAGGCGCC 34358
QY 2501 CTGGGGATGCCCAAGCGCAGAAATACGGAACACAGAAATCAGGGAGAAATCATGTGA 2559
Db 34359 CTGGGGATGCCCAAGCGCAGAAATACGGAACACAGAAATCAGGGAGAAATCATGTGA 34417

RESULT 9
LOCUS AX417281 8001 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 3 from Patent WO0230981.
ACCESSION AX417281
VERSION AX417281.1 GI:21522590
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
DIFFERENCECO.V., Beasley,E.M., Ketchum,K.A., Wei,M.-H. and Zhong,W.
TITLE Gene encoding a human g-protein coupled receptor and its use
JOURNAL Patent: WO 0230981-A 3 18-Apr-2002;
PE CORP NY (US)
FEATURES
Source 1..8001
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1291 a 2632 c 2623 g 1455 t
ORIGIN
Query Match 76.1%; Score 1947.4; DB 6; Length 8001;
Best Local Similarity 82.0%; Pred. No. 4,3e-251;
Matches 2558; Conservative 0; Mismatches 1; Indels 560; Gaps 5;
QY 1 ATGCTGGGCGCTGTGCTGTGCGCTCAGCCTGTGGGCTCTCTGCAACCTGGAGCGGG 60
Db 2118 ATGCTGGGCGCTGTGCTGTGCGCTCAGCCTGTGGGCTCTCTGCAACCTGGAGCGGG 2177
QY 61 GCCCATTTGTCCTGTACAGCACTTAAGATGAAGGGGACTACGTGTGTGGGGGCTG 120
Db 2178 GCCCATTTGTCCTGTACAGCACTTAAGATGAAGGGGACTACGTGTGTGGGGGCTG 2237
QY 121 TTCCCTGGGCGAGGCGAGGAGGCTGCGCTCCGACGCGAGACGCGCCAGAGGCT 180
Db 2238 TTCCCTGGGCGAGGCGAGGAGGCTGCGCTCCGACGCGAGACGCGCCAGAGGCT 2297
QY 181 GTGTGCA-----187
Db 2298 GTGTGCAAGGTACAGAGGTGGGAGCGGCTCGGGGTACAGGTGACAGGTCTGGG 2357
QY 188 -----CCAGGTTCCT 199

```

```

Db 2358 GTGCTCTGAGTGGGCGCAGGTGCGCATCTGCGGTTCTGTGTGGCCAGGTTCTCT 1111111111
QY 200 CAAAGGCGCTGTGCGCACTGGGCAATGAAGATGCGCGGAGAGATCAACAAAGT 259
Db 2418 CAAAGGCGCTGTGCGCACTGGGCAATGAAGATGCGCGGAGAGATCAACAAAGT 2477
QY 260 CGGATCTGTGCGCGGCGCTGCGCTGCGCTACGACCTCTTGTGATAGTGTCTGAGGCTG 319
Db 2478 CGGATCTGTGCGCGGCGCTGCGCTGCGCTACGACCTCTTGTGATAGTGTCTGAGGCTG 2597
QY 320 TGTGTGCGCATGAAGCCAGCCTATGTTCTGTGCGCAAGGAGGAGCGGACATCGCG 379
Db 2538 TGTGTGCGCATGAAGCCAGCCTATGTTCTGTGCGCAAGGAGGAGCGGACATCGCG 2597
QY 380 CTTACTGCAATACAGCAGTACAGGCGCGGCTGCGGCTGCGGCTGATGAGGCGCGCTGCT 439
Db 2598 CTTACTGCAATACAGCAGTACAGGCGCGGCTGCGGCTGCGGCTGATGAGGCGCGCTGCT 2657
QY 440 CAGAGCTCGCATGTGTCACGCGGCAAGTTCTTACGCTTCTTCTCATGCC-----489
Db 2658 CAGAGCTCGCATGTGTCACGCGGCAAGTTCTTACGCTTCTTCTCATGCCAGGTTGGCG 2717
QY 490 -----489
Db 2718 CCCCCACATCAACCCACCCCGAGCCCTGCGCGGAGGCCCTGTGTACAGAGAT 2777
QY 490 -----CAGTCAGCTACGTGTGTAGATGAGGCTGTGAGCGCGCGCGGAG 534
Db 2778 GCGCTTGGCGCTGTGCAAGTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2837
QY 535 ACCTTCCCTCTCTTCTTCCGACCGCTGCCAGCGACCGCTGTGACGTGACGCGCGCG 594
Db 2838 ACCTTCCCTCTCTTCTTCCGACCGCTGCCAGCGACCGCTGTGACGTGACGCGCGCG 2897
QY 595 GAGCTGTGAGAGAGTTCGCTGGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 654
Db 2898 GAGCTGTGAGAGAGTTCGCTGGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2957
QY 655 GGCGGCGAGGCGCTGAGCATCTTCTCGGCCCTGCGCGCGCGCGCGCGCATCTGATCGCG 714
Db 2958 GGCGGCGAGGCGCTGAGCATCTTCTCGGCCCTGCGCGCGCGCGCGCGCATCTGATCGCG 3017
QY 715 CACGAGGCGCTGTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 774
Db 3018 CACGAGGCGCTGTGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3077
QY 775 GTCTGTACAGAGTGAACAGAGAGCGGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 834
Db 3078 GTCTGTACAGAGTGAACAGAGAGCGGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3137
QY 835 GCGCGCCAGCGCTCTTCAACTACAGCATAGCAGCAGCGCTTCCGCCAAGGTGTGGTG 894
Db 3138 GCGCGCCAGCGCTCTTCAACTACAGCATAGCAGCAGCGCTTCCGCCAAGGTGTGGTG 3197
QY 895 GCCAGCGAGGCGCTGTGACCTGTGACCTGTGATGAGGCGCTGCCCGCATGCGCAATG 954
Db 3198 GCCAGCGAGGCGCTGTGACCTGTGACCTGTGATGAGGCGCTGCCCGCATGCGCAATG 3257
QY 955 GCGAGGCTGTGTGCTTCTTCCAGAGGCGTCCAGGTGTGACAGGTTCCCGAGTAGTG 1014
Db 3258 GCGAGGCTGTGTGCTTCTTCCAGAGGCGTCCAGGTGTGACAGGTTCCCGAGTAGTG 3317
QY 1015 AAGAGCGACCGCGCGCTGCGCGACGACCGCGCTTGTGCTGTGCGCGCGGAGAGGAG 1074
Db 3318 AAGAGCGACCGCGCGCTGCGCGACGACCGCGCTTGTGCTGTGCGCGCGGAGAGGAG 3377
QY 1075 CAGGCTGTGAGAGAGAGTGTGTGCGCAGCGCTGCCGAGTGTGACTCATACGCTG 1134
Db 3378 CAGGCTGTGAGAGAGAGTGTGTGCGCAGCGCTGCCGAGTGTGACTCATACGCTG 3437
QY 1135 CAGAGCTGAGCGAGGCGTAAATCACACACAGAGCTTCTGTGTGTGTGTGTGTGTGTGT 1194

```

Db 3438 CAGAACGTGAGCGGAGGCTAAATCATCACACAGACGTTCTGTCTACGACGCTGTGTAT 3497
QY 1195 AAGGTGGCCAGGCGCTGTGACAAACACTTCTGATGCAAGGCTCAGAGCTGCCCGGCGAG 1254
Db 3498 AAGGTGGCCAGGCGCTGTGACAAACACTTCTGATGCAAGGCTCAGAGCTGCCCGGCGAG 3557
QY 1255 GACCCCGTGAAGCCCTG----- 1271
Db 3558 GACCCCGTGAAGCCCTGAGGAGGTGAGCCCGGAGATGGGGGTGTGCTGTCTGTGATGT 3617
QY 1272 ----- 1271
Db 3618 GCCCAGGCGCACAGGACGCGCACACAGCCTGAGTGTGATGTGCGGCTCAGCCCG 3677
QY 1272 -----GAGCTCTGTGAGAACATGTATCAACCTTACACCTTGGGGGCGGCTG 1320
Db 3678 GTCCCCCGCGGAGCTCTGTGAGAACATGTATCAACCTTACACCTTGGGGGCGGCTG 3737
QY 1321 CGGCTGCGGTTGACAGCAGCGGAAAGTGTGACATGATGATGACACTTGAAGCTGTGGTG 1380
Db 3738 CGGCTGCGGTTGACAGCAGCGGAAAGTGTGACATGATGATGACACTTGAAGCTGTGGTG 3797
QY 1381 TGGCAGGCTCAGTGTGCGGCTCCAGCAGTGTGGGAGGTTCAACGCGACCTCAGAGACA 1440
Db 3798 TGGCAGGCTCAGTGTGCGGCTCCAGCAGTGTGGGAGGTTCAACGCGACCTCAGAGACA 3857
QY 1441 GAGCGCTGAAGATCGGCTGGGAGCAGCAGTGTGACAAAC----- 1476
Db 3858 GAGCGCTGAAGATCGGCTGGGAGCAGCAGTGTGACAAACAGGTTAGTGAAGGTGGGTGTGC 3917
QY 1477 ----- 1476
Db 3918 CAGGCGTGCCTGTGTAGCCCGCGGCGAGAGGCGCAGCCTGGGGGCTGGGGCGCTTCAG 3977
QY 1477 -----CAGAACCCG 1486
Db 3978 TCTCCCGTGGGAGATGCGCAGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4037
QY 1487 TGTCCCGGTGTGTGCGGCGAGTGTGCGAGAGGCGCAGGTTGGCGGGGTTCCACT 1546
Db 4038 TGTCCCGGTGTGTGCGGCGAGTGTGCGAGAGGCGCAGGTTGGCGGGGTTCCACT 4097
QY 1547 CCTGTGTGACAGTGTGTGAGCTGTGAGCGGCGGAGCTACCGGCAAAAC----- 1597
Db 4098 CCTGTGTGACAGTGTGTGAGCTGTGAGCGGCGGAGCTACCGGCAAAACCGAGTGAGC 4157
QY 1598 ----- 1597
Db 4158 CGCCTTCCGCGGAGCGGCGGAGTGTGGAACGACAGAGGAGGCTTCGCCAAGTCTGACT 4217
QY 1598 -----CAG 1600
Db 4218 CTGAGACGAGAGCCAGGAGGAGCAAGAGAACACACAGCGCCCTTCTCTCTACAG 4277
QY 1601 ACGACATGCGCTGTGACCTTTTGTGGCAGAGATGATGTCCCGGAGCAGAACACAGCT 1660
Db 4278 ACGACATGCGCTGTGACCTTTTGTGGCAGAGATGATGTCCCGGAGCAGAACACAGCT 4337
QY 1661 GCTTCCGCGGAGAGTGTGCTTCTGTGAGTGTGGGCGAGCCGGCTGTGTGTGTGTCTCC 1720
Db 4338 GCTTCCGCGGAGAGTGTGCTTCTGTGAGTGTGGGCGAGCCGGCTGTGTGTGTGTCTCC 4397
QY 1721 TGTGTGTGAGAGCTGGGCGCTGT 1780
Db 4398 TGTGTGTGAGAGCTGGGCGCTGT 4457
QY 1781 GGGACAGCCCACTGTGTGAGGCGCTGTGGGCGGCGCCCTGTGCTGTGTGTGTGTGTGTGT 1840
Db 4458 GGGACAGCCCACTGTGTGAGGCGCTGTGGGCGGCGCCCTGTGCTGTGTGTGTGTGTGTGT 4517
QY 1841 TGGGCGT 1900
Db 4518 TGGGCGT 4577

QY 1901 TGGCCAGAGAGCGCTGT 1960
Db 4578 TGGCCAGAGAGCGCTGT 4637
QY 1961 AGCGGCGGAGATCTTGTGTGAGTGTGAGAACCTGCTGTGAGTGTGGGAGCAGCCGCTGAGTG 2020
Db 4638 AGCGGCGGAGATCTTGTGTGAGTGTGAGAACCTGCTGTGAGTGTGGGAGCAGCCGCTGAGTG 4697
QY 2021 GCTGCTGTGGGGGCGCTGTGGGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2080
Db 4698 GCTGCTGTGGGGGCGCTGTGGGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4757
QY 2081 CACTGTGACCTGT 2140
Db 4758 CACTGTGACCTGT 4817
QY 2141 TGGCCAGAGAGCGCTGTGTGACTGTGCGGACAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2200
Db 4818 TGGCCAGAGAGCGCTGTGTGACTGTGCGGACAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4877
QY 2201 AGCCCAACAAATGCGACGCTGTGCTTGT 2260
Db 4878 AGCCCAACAAATGCGACGCTGTGCTTGT 4937
QY 2261 AGCCGCGGCTGT 2320
Db 4938 AGCCGCGGCTGT 4997
QY 2321 CTTGGGT 2380
Db 4998 CTTGGGT 5057
QY 2381 AGATGGGCGGCTGT 2440
Db 5058 AGATGGGCGGCTGT 5117
QY 2441 GTTACTGT 2500
Db 5118 GTTACTGT 5177
QY 2501 CTGGGGATGCCCAAGGCGCAGATGACGGGAGCAACAGAGAAATGAGGGGAAACATGAGTGA 2559
Db 5178 CTGGGGATGCCCAAGGCGCAGATGACGGGAGCAACAGAGAAATGAGGGGAAACATGAGTGA 5236

RESULT 10
AC026283 197748 bp DNA linear HTG 07-JUL-2000
LOCUS Homo sapiens chromosome 1 clone RP11-421C4, WORKING DRAFT SEQUENCE,
DEFINITION 17 unordered pieces.
ACCESSION AC026283.2 GI:523967
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 197748)
AUTHORS Waterston, R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 197748)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Apr 7, 2000 this sequence version replaced gi:7272307.

	misc_feature	/note="assembly_name:Contig7"	9402.	12779	
	misc_feature	/note="assembly_name:Contig8"	12880.	15898	
	misc_feature	/note="assembly_name:Contig9	clone_end:"f7		
	misc_feature	vector_side:right"	15999.	.21627	
	misc_feature	/note="assembly_name:Contig10"	21728.	.28379	
	misc_feature	/note="assembly_name:Contig11"	28480.	.36334	
	misc_feature	/note="assembly_name:Contig12"	36435.	.50851	
	misc_feature	/note="assembly_name:Contig13"	50952.	.60828	
	misc_feature	/note="assembly_name:Contig14"	60929.	.71240	
	misc_feature	/note="assembly_name:Contig15"	71341.	.88918	
	misc_feature	/note="assembly_name:Contig16"	89019.	.108716	
	misc_feature	/note="assembly_name:Contig17"	108817.	.128174	
	misc_feature	/note="assembly_name:Contig18"	128275.	.164210	
	misc_feature	/note="assembly_name:Contig19"	164311.	.197748	
	misc_feature	/note="assembly_name:Contig20"			
BASE COUNT	42194 a	54729 c	55209 g	44011 t	1605 others
ORIGIN					
Query Match	75.2%	Score 1923.8;	DB 2:	Length 197748;	
Best Local Similarity	82.0%;	Pred. No. 3e-248;			
Matches 2555;	Conservative	0; Mismatches	2; Indels	560; Gaps	7;
OY	1	ATGCTGGGCCCTCGCTGTCCCTGAGCCTCTGGGCTCTCTGCACACCAGGAGCGGG	60		
Db	71768	ATGCTGGGCCCTCGCTGTCCCTGAGCCTCTGGGCTCTCTGCACACCAGGAGCGGG	71827		
OY	61	GCCCCATTGTGCTGTCAACAGCACTTATGATGAAGGGGACTACGTCTGGGGGGCTG	120		
Db	71828	GCCCCATTGTGCTGTCAACAGCACTTATGATGAAGGGGACTACGTCTGGGGGGCTG	71887		
OY	121	TTCCTCCCTGGGGGAGGCGGAGAGGGCTCGCGACCGGACACGGCCACAGACCCCT	180		
Db	71888	TTCCTCCCTGGGGGAGGCGGAGAGGGCTCGCGACCGGACACGGCCACAGACCCCT	71947		
OY	181	GTGTGCA-----	187		
Db	71948	GTGTGCA--CAGGTTCCTCT	72007		
OY	188	-----CAGGTTCCTCT	199		
Db	72008	GTGCTCTGAGCTGGGGGCCAGGTGGCCATCTGGCGTTCTGTGTGGGCCCAAGTTCTCT	72067		
OY	200	CAAACGAGCTGTGCTGGGACTGEGCCATGAATAATGGCCGTGGAGAGATCAACAAGA	259		
Db	72068	CAAACGAGCTGTGCTGGGACTGEGCCATGAATAATGGCCGTGGAGAGATCAACAAGA	72127		
OY	260	CGGATCTGCTGCGCGGGCTGGCGGCTTGAGATCATGCTGTGCTGTGCTGTGCTGTG	319		
Db	72128	CGGATCTGCTGCGCGGGCTGGCGGCTTGAGATCATGCTGTGCTGTGCTGTGCTGTG	72187		
OY	320	TGGTGGCATGAAGCCAGCCCATGATGTCCTGAGGCAAGGAGGAGCGGCGCATCGCG	379		
Db	72188	TGGTGGCATGAAGCCAGCCCATGATGTCCTGAGGCAAGGAGGAGCGGCGCATCGCG	72247		
OY	380	CCTACTGCACTACACGAGTACAGCCCGGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	439		
Db	72248	CCTACTGCACTACACGAGTACAGCCCGGTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	72307		
OY	440	CAGAGCTGCGCATGAGTCAACGAGCAATGTTTCAGCTTCTTCCTCATGAGCC-----	489		

QY 2203 GCCACCAATGCCAGCTGGCCCTTCTGCTTCTGCTGGGCACTTCTGCTGGAGCCAG 2262
Db 74526 GCCACCAATGCCAGCTGGCCCTTCTGCTTCTGCTGGGCACTTCTGCTGGAGCCAG 74585
QY 2263 CCGGGCTCTCAACACCGCTGGCTGCTCACTTCTGCTGGAGCTTCTGCTGGAGCC 2322
Db 74586 CCGGGCTCTCAACACCGCTGGCTGCTCACTTCTGCTGGAGCTTCTGCTGGAGCC 74645
QY 2323 TGGGCTCTTCTGCTGGCCCTCTGCTGGCAATGCTGCTGCTGGAGCCAGCTGGAG 2382
Db 74646 TGGGCTCTTCTGCTGGCCCTCTGCTGGCAATGCTGCTGCTGGAGCCAGCTGGAG 74705
QY 2383 ATGGGCGCCCTCTGCTGCTGCTGCTGGCACTTCTGCTGGAGCTTCTGCTGGAG 2442
Db 74706 ATGGGCGCCCTCTGCTGCTGCTGCTGGCACTTCTGCTGGAGCTTCTGCTGGAG 74765
QY 2443 TACCTGCTCATGCGGCGCAGGCTCAACACCCCGAGTCTTCTGCTGGAGGCGCT 2502
Db 74766 TACCTGCTCATGCGGCGCAGGCTCAACACCCCGAGTCTTCTGCTGGAGGCGCT 74825
QY 2503 GGGGATGCGCCAGGCGCAATGAGGAGACAGGAAATCAAGGGAACATGAGTGA 2559
Db 74826 GGGGATGCGCCAGGCGCAATGAGGAGACAGGAAATCAAGGGAACATGAGTGA 74882

RESULT 11
AX451560 1788 bp DNA linear PAT 03-JUL-2002
LOCUS AX451560
DEFINITION Sequence 4 from Patent WO0224885.
VERSION AX451560.1 GI:21698546
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1
AUTHORS Kossida, S.
TITLE Regulation of human g protein-coupled receptor
JOURNAL Patent: WO 0224885-A 4 28-MAR-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
Source 1. 1788
Location/Qualifiers

BASE COUNT 236 a 620 c 593 g 339 t
ORIGIN

Query Match 61.5%; Score 1574; DB 6; Length 1788;
Best Local Similarity 89.8%; Pred. No. 4.2e-201;
Matches 1788; Conservative 0; Mismatches 0; Indels 204; Gaps 1;

QY 490 CAGGTACAGTACGCTGCTAGATGAGTGTGCTGAGCGCCCGGAGACCTTCCCTCTTC 549
Db 1 CAGGTACAGTACGCTGCTAGATGAGTGTGCTGAGCGCCCGGAGACCTTCCCTCTTC 60
QY 550 TTCCGCAACGCGCCAGGAGACCTGTGAGCTGAGCGCGCGGAGAGCTGCTGAGAG 609
Db 61 TTCCGCAACGCGCCAGGAGACCTGTGAGCTGAGCGCGCGGAGAGCTGCTGAGAG 120
QY 610 TTCCGCTGGAAGTGGTGGCCCTTGGGACGAGCAAGTACGAGCGCGGAGAGCTG 669
Db 121 TTCCGCTGGAAGTGGTGGCCCTTGGGACGAGCAAGTACGAGCGCGGAGAGCTG 180
QY 670 AGCATCTTCTGCGCCCTGCGCGGCGCAGCGGCACTTCTGATCGCGCAGAGGCTGTG 729
Db 181 AGCATCTTCTGCGCCCTGCGCGGCGCAGCGGCACTTCTGATCGCGCAGAGGCTGTG 240
QY 730 CCGGTGCGCGCTGCGGATGACTGCGGCTGGGAGAGTGAAGAGCTGCTGACAGAGT 789
Db 241 CCGGTGCGCGCTGCGGATGACTGCGGCTGGGAGAGTGAAGAGCTGCTGACAGAGT 300

QY 790 AACCAAGACAGCTGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 849
Db 301 AACCAAGACAGCTGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 850 TTCAACTACACATCAGCAGAGCTCTGCGCCAAAGTGTGGTGGCCAGAGAGGCTGTG 909
Db 361 TTCAACTACACATCAGCAGAGCTCTGCGCCAAAGTGTGGTGGCCAGAGAGGCTGTG 420
QY 910 CTGACCTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 969
Db 421 CTGACCTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 970 TTCTCCAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1029
Db 481 TTCTCCAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 1030 CTGGCCACCGACCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1089
Db 541 CTGGCCACCGACCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 1090 GAGGTGTGGGCGACGCGCTGCGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149
Db 601 GAGGTGTGGGCGACGCGCTGCGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 1150 GGGCTAATATCACACACAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1209
Db 661 GGGCTAATATCACACACAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 1210 CTGCACAACACTTTCAGTGTGCAAGCGCTGAGGCTGCGCGGAGAGACCCCGTAA 1269
Db 721 CTGCACAACACTTTCAGTGTGCAAGCGCTGAGGCTGCGCGGAGAGACCCCGTAA 780
QY 1270 TGGCAGCTCTGAGAAACATATACACTTTCACGCTGAGGCGGCGCTGCGCTGCG 1329
Db 781 TGG----- 783
QY 1330 TTGCACAGACAGCGGAACGTTGAGATGAGTACAGCTGAAAGCTGTGGTGTGCGAG 1389
Db 784 ----- 783
QY 1390 TCAGTGGCCAGGCTCCACAGAGTGTGGGAGTTTCAAGCGCACCTCAGAGAGCGCT 1449
Db 784 ----- 783
QY 1450 AAGATCCGCTGGCACAGCTCTGACACCGAAGCCCGTCCCGTGTGCTGCGGAGTGC 1509
Db 784 -----CAGAAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 816
QY 1510 CAGAGAGGCGCAGGTGCGCGCGGCTCAAGGGGTTTCACTCTGCTGCTGCTGCTGCT 1569
Db 817 CAGAGAGGCGCAGGTGCGCGCGGCTCAAGGGGTTTCACTCTGCTGCTGCTGCTGCTGCT 876
QY 1570 TGGCAGGCGGCGACCTTCCGCAAAACCGAAGCAACATGCTGCTGCTGCTGCTGCT 1629
Db 877 TGGCAGGCGGCGACCTTCCGCAAAACCGAAGCAACATGCTGCTGCTGCTGCTGCTGCT 936
QY 1630 GATGAGTGTGCTGCGGAGCAGACAGAGCTGCTGCGCCAGAGTCTGCTGCTGCTGCT 1689
Db 937 GATGAGTGTGCTGCGGAGCAGACAGAGCTGCTGCGCCAGAGTCTGCTGCTGCTGCT 996
QY 1690 TGGGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1749
Db 997 TGGGCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1056
QY 1750 CTGGCTGCTTGGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1809
Db 1057 CTGGCTGCTTGGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1116
QY 1810 GGGGCGCTGCGCTCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1869
Db 1117 GGGGCGCTGCGCTCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1176
QY 1870 TTCCCTGCGCAGCCAGCCCTGCGGATGCTGCGGCGCAGAGCCCTTGTCCACCTCCG 1929

Db 541 CCATCTCTCTCCGACAGTCCAGAGTACCGGGTGACAGTGCAGGCCGTTGTGACACTG 600
 QY 601 CTGAGAGAGTTCGGCTGGAAGTGGTGGCCGCCCTGGGACACGACGACGATGAGCCGG 660
 Db 601 TTGCGAATTCAGTGGAGTGGAGTGGCTCTTGGATGATGATGATGATGATGATGATG 660
 QY 661 CAGGGCCCTGACATCTTCTCGGCCCTGGCCCGGACGCGGACATCTGATCGCGACAG 720
 Db 661 GAAGGTCTGACATCTTCTTCTGCTGGCCAACTCAGAGGATCTGATTCAGACAGAG 720
 QY 721 GGCCTGGCCGCTGCCCGGCGGATGATCGGCTGGGGAGAGTGTGAGAGCTCTG 780
 Db 721 GGCCTGGGCGACACATGACATGATGACCAATTTGGGACAGTGTGATGTGCTA 780
 QY 781 CACAGGTGAACAGAGAGGCTGAGGTGGTGGCTGTTCGCTGCGCTGCTGACAGCCGC 840
 Db 781 CGCCAGTGAACAGAGAGGATGAGGTGGTGGCTGTTCGCTGCGCTGCTGACAGCCGC 840
 QY 841 CAGGCCCTCTTCAACTACAGATGACAGAGGCTCTGCGCCAAAGTGTGGTGGCCAGC 900
 Db 841 TACTCCCTTTTACGATACAGATGACATGACCTCTGACCCAAAGGTATGGGTGGCCAGT 900
 QY 901 GAGGCTGTGAGTCTGACCTGACCTGATGATGGGCTGCGCGCATGCGCCAGATGGCCAG 960
 Db 901 GAGTCTGTGAGTCTGACCTGACCTGATGATGGGCTGCGCGCATGCGCCAGATGGCCAGT 960
 QY 961 GTGCTGGCTCTCTCAGAGAGGGTGGCCAGTGGACAGAGTTCGCCAGTATGATGAGAG 1020
 Db 961 GTTCTTGGGTTTCTGACAGCGGGTGGCCAGTGGACAGAGTTCGCCAGTATGATGAGAGT 1020
 QY 1021 CACCTGGCCCTGAGCAGCAGCAGCCGCTTCTGCTCTGCTGCGCGGAGAGAGAGAGAGT 1080
 Db 1021 CGCCTGGCCCTGAGCAGCAGCAGCCGCTTCTGCTCTGCTGCGCGGAGAGAGAGAGT 1080
 QY 1081 CTGAG 1140
 Db 1081 CTGAG 1140
 QY 1141 GTGAGCGAGAGAGCTAA-----ATCACACACAGAGCTTCT 1176
 Db 1138 CTGATCATCTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197
 QY 1177 GTCTACGAGAGTGTGTATAGGCTGGCCAGGCTGCGCAACACTCTTCACTGATGAGAGCC 1236
 Db 1198 ACCTATGAGAGTGTGTATAGGCTGGCCAGGCTGCGCAACACTCTTCACTGATGAGAGCC 1236
 QY 1237 TCAG 1296
 Db 1258 TCACATGAG 1317
 QY 1297 CTGACCTTCACTGAG 1356
 Db 1318 ATGATGTTCCGCTGCTGAGACTTTCGACTGATGATGATGATGATGATGATGATGATG 1377
 QY 1357 GAGTACAGAGT 1416
 Db 1378 GAATATGAG 1437
 QY 1417 AGGTTCAAG 1476
 Db 1438 ACCTTCAAG 1491
 QY 1477 CAGAAGCCCGTGTCCCGT 1536
 Db 1492 CAGGTGCGAGTGTCCAGT 1551
 QY 1537 GGGTTCACCTCTGT 1596
 Db 1552 GGGTTCATCTCTGT 1611
 QY 1597 CCAG 1656
 Db 1612 CCAATGATCTTCACTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 1671

QY 1657 CGTCTTCGCGCGAGAGTCTGGTTCCTGGCATGGGGCGAGCCGGCTGTGCTGTGCTG 1716
 Db 1672 ACCTGTACTCTGAGAGCCCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1731
 QY 1717 CTCTGTCTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1776
 Db 1732 CTCTGTCTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1791
 QY 1777 CATGGGAG 1836
 Db 1792 TACTGGAG 1851
 QY 1837 TGCCGTGGAG 1896
 Db 1852 TGCCGTGGAG 1911
 QY 1897 TGCCGTGGAG 1956
 Db 1912 TGCCGTGGAG 1971
 QY 1957 CTGAG 2016
 Db 1972 CTGAG 2031
 QY 2017 AGTGGCTGCTGCGGGGAG 2076
 Db 2032 TGACAGTCTCTGCGGGGAG 2091
 QY 2077 GTGCACTGTGACCTGT 2136
 Db 2092 GTGCACTGTGACCTGT 2151
 QY 2137 ATGCTGCCAG 2196
 Db 2152 GTGCTGCCAG 2211
 QY 2197 GCGCAGCGACCAATGCGACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2256
 Db 2212 GTGCACTGTGACCTGT 2271
 QY 2257 AGCCAGCGGGGCTGT 2316
 Db 2272 AGCCAGCGGGGCTGT 2331
 QY 2317 ATCACTGT 2376
 Db 2332 ATCACTGT 2391
 QY 2377 GTGAGATGGGCGGCTGT 2436
 Db 2392 GTGAGATGGGCGGCTGT 2451
 QY 2437 AGT 2496
 Db 2452 AAATGT 2511
 QY 2497 GGCCCTGGGAGTGGCCAAAGGCGCAAGATGACGGGA 2530
 Db 2512 AGCCCAAGAGAGATGATGGGAATGTGTGTA 2545

RESULT 13
 AY032620 2577 bp mRNA 11near ROD 05-FEB-2002
 LOCUS
 DEFINITION Rattus norvegicus sweet taste receptor T1R3 mRNA, complete cds.
 ACCESSION AY032620
 VERSION AY032620.1 GI:15147674
 KEYWORDS
 SOURCE Rattus norvegicus.
 ORGANISM Rattus norvegicus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

QY	601	CAAGAGAGTTGGCGTGGAACTGGGGGGCCCGCTGGGACAGACAGCACTACGGCCG	660
Db	601	TTGGAGAACTTACGCTGGAACTGGGGGCCGCTTAGGAGAGTAAGTAAGTAAGTACG	660
QY	661	CAGGGCCGTGAGCAATCTTCGCGCCCTGGCGGCGGACAGCGGGCATCTGCATCGCCACAG	720
Db	661	GAAGGTCTGAGCAATCTTCTTACTGCTGGCCAAATGCACAGGATCTGCATCGCAATGAG	720
QY	721	GGCTTGGTGGCCGCTGCCCGGTGGCCATGACTGCGCGCTGGGAGAGTGTACAGACGTC	780
Db	721	GGCTTGGTGGCCACCAACATGCACTAAGTGGCCCAAGTTGGGGAAGAGTGTGGATGTACTA	780
QY	781	CACCAAGTGAACACAGACAGACGCTGACAGTGGTGTGCTGTGCTGCGCTCAGCGCCG	840
Db	781	CGCCAAATGAACCAAGTAAGTAAGTAAGTAAGTGAAGTGAAGTAAGTAAGTGAAGTGA	840
QY	841	CAGCGCCTCTTCAACTACATACAGATCAGACAGAGCTGTGCGCCAAAGTGTGGTGGCAGC	900
Db	841	TACTCCCTTTTGTAGTACAGACTCCATCATGTGGCTCTCCACCAAGTATGGGTGGCAGT	900
QY	901	GAGGCGCTGGTGACCTGTACCTGGTACAGGGGGTGGCGCGCATGGCCCAATGGGACG	960
Db	901	GAGTCTTGGCTGTACATCTGACCTGGTACAGACTTCCCAATATATGGTGGGACCT	960
QY	961	GTGCTTGGCTTCTCCAGAGGGGTGGCCAGCTGACAGAGTTCCCCCACTACGTAAAGC	1020
Db	961	GTGCTTGGTGTTTTGGACGGGGGTGCCCTACGTCCGTAATTTTCCATATATGTGAGACT	1020
QY	1021	CACCTGGCCCTGGGACCAACCCGCGCTTCTGCTGTGCCCTGTGGCGAGAGGACAGGAT	1080
Db	1021	CACCTTGCCTGGCGCCTGACCCAGCATTTCTGTGCTCACTG--AAAGCGAATTTGGAT	1077
QY	1081	CTGAGAGAGACCTGGTGGGACAGCGCTCCGCGACAGTGTGACTGATCAGCC-----	1132
Db	1078	CTGAGGAGAAATGTGATGGGGCAACGCTGTCCACGCTGTGACAGCATATCTGTGACAAAC	1137
QY	1133	-----TGCAGAACTGAGCGCAGGGCTTAATCACCACCAAGCTTCTCT	1176
Db	1138	CTATCATCTGGGCTGTGTGAGAAACCTATGACGTGGGCAATTGCAACCAATATTTTGCA	1197
QY	1177	GTCATACCACTGTGTATATAGCGGGCCAGGCGCTGCAACAACATCTTCAATGGACAAGC	1236
Db	1198	ACGTATGACACTGTGTACAGTGTGTGCTCAAGCCCTTTCACAAACCCCTACATGTGCAATGTC	1237
QY	1237	TCAGGCTGCCCGCGCAGGAGACCCCGTGAAGCCCTGGCAGACTCTGGAGAACATGTACAC	1296
Db	1258	TCACATTTGCCACGTATCAGAACATATGTTCTTACCTGTGGACGCTCTGTGAACATGTACAAAT	1317
QY	1297	CTGACCTTCCACGTGGGGCGGCGTGGCGGTGCGAGTGTGACAGAGGAGGAAGTGTGACATG	1356
Db	1318	ATGAGTTTCCATGTCTGAGACTTGTACACTACAGATTGTGATGCTGGAAGGGAATGTAGACATG	1377
QY	1357	GAGTACGACCTGAAGCTGTGTGGTGTGGCAGGGCTCAGTGTGCCAGGCTCCACAGAGTGGGC	1416
Db	1378	GAATATATGACCTGAAGATGTGGGTGTGGCAGAGCCGTACACCTGTATATATATGTGGCG	1437
QY	1417	AGGTTCACCGCAGCCTCAGAGACAGAGCGCCCTGAAGATCCGTGGCAACAAGTGTGACAAAC	1476
Db	1438	ACCTTTCACGCGCACCTCTTACGTGACGTGCAAGTCTAAATAATGTACTTGGC-----CAGCGAC	1491
QY	1477	CAGAAAGCCCTGTCCGCTCTGTGCGGCGAGTGCAAGAGAGGCGCAGTGTGCGCGGGTCAAG	1536
Db	1492	CAGGTGGCACTGTCCCAAGTGTTCGCCCGCAGTGCMAAAGATGGCCAGGTGTGCCGAGTAAG	1551
QY	1537	GGGTTCACCTCTGTCTACGACTGTGTGACCTGCGAAGTGGCGGCGAGCTTACCGGCAAAAC	1596
Db	1552	GGCTTTCATTCCTGTGCTGTATGTACTCCTGTGGACTCAGAGCGCGCGACGTACCGGAAGCAT	1611
QY	1597	CCAACAGCAATCGCGTGCACCTTTTGTGGCAGGAATGATGTGTCCCGGAGAGGAAGACA	1656
Db	1612	CCAATATACTTCACTCTTACTCCATGTATTAACCAAGACCAATGTGTCCCGCAGAAAGACACA	1671
QY	1657	CGCTGTTCGCGCGCAGGTCGTGGTTCGTGGCATGGGGGAGACCGCGTGTGCTGTGCTG	1716

[illegible]

AUTHORS Bachmanov, A.A., Li, X., Reed, D.R., Ohmen, J.D., Li, S., Chen, Z., Tordoff, M.G., deJong, P., Wu, C., West, D.B., Chatterjee, A., Ross, D.A. and Beauchamp, G.K.

TITLE Positional cloning of the mouse saccharin preference (Sac) locus

JOURNAL Chem. Senses 26 (7), 925-933 (2001)

MEDLINE 21439023

PubMed 1155487

REFERENCE 2 (bases 1 to 2577)

AUTHORS Li, X., Bachmanov, A.A., Li, S., Chen, Z., Tordoff, M.G., Beauchamp, G.K., deJong, P., Wu, C., Chen, L., West, D.B., Ross, D.A., Ohmen, J.D. and Reed, D.R.

TITLE Genetic, physical, and comparative map of the subtelomeric region of mouse Chromosome 4

JOURNAL Mamm. Genome 13 (1), 5-19 (2002)

MEDLINE 21635547

PubMed 11773963

REFERENCE 3 (bases 1 to 2577)

AUTHORS Bachmanov, A.A., Li, X., Reed, D.R., Ohmen, J.D., Li, S., Chen, Z., Tordoff, M.G., deJong, P., Wu, C., West, D.B., Chatterjee, A., Ross, D.A. and Beauchamp, G.K.

TITLE Direct Submission

JOURNAL Submitted (05-OCT-2000) Monell Chemical Senses Center, 3500 Market Street, Philadelphia, PA 19104, USA

FEATURES

source

1..2577

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/chromosome="4"

/map="83.0cM"

1..2577

/gene="Tas1r3"

/allele="Sac-b"

1..2577

/gene="Tas1r3"

/codon_start=1

/product="putative sweet taste receptor type 1 member 3"

/protein_id="AL08425.1"

/db_xref="GI:15808695"

/translation="MPALAIIMGLSLAFLLELGMASISLCSQKFAAGDYILGLFLPIG STEEATLNOETOPNSTPCNRFSPILGLFLMAMKAVEELINNSALLGLRLGLDPT CSRPVYTKMSLFLAKVGSOSIAACNTQYQPRVLAIVGRHSSELALITKRFSEF LMPQVTSASMDRLSDRETFPSFRIVPSDRVOLAIVLILQFNMKNVAULSDDY GRPQVTSFSLANARIGICIAHEGLVPOHDTSGOLKIVLDVLAQVMSKVQVVLAS ARAYVSLFYSIHGSLSPKVMVASBSMLSDVLPNLAIVLAVGVGLFIRGLALFEF SHYVETHLIALAPARCASLANLELDEEHMGRCRCDIMONTSSGLIOLWSAGO LHHOIRATYAAVSVNAOLAHNTLOCNVSCHVSEHVLPMOLLNNMNSPHADLTIO FPAEGVNDMEYDLKMWOSPTVLTHTVGTFTNLTLOOSKATPFGNOYVSCSRQC KDQVRRVGFHSCCYDCVCKAGSYRKHDDFTCPNDDWSPEKSTACLERRPF LAMGEVNVLSLALICLVGLALALGLSVHMDSPLYVAGSGSQCFGLICGLFCL SVLLFPGRRSSASCLAOQPMALPLTGCSTLFLQAAFEVSELPISMANMSTYLR GLAMLVVILATVEALCAWYLIAPPEVDTGVALEFVLEHCHVRSVSLGLYR TNAMLAFLCFLGFEIYLOSQGRYNRARGLTFMALVFTVWSFPLIANOVYQYAV QKGAIIYVCAIGIIVFTHLPRCYULMLPLINTDFELGRNAKAADENSQGGAAAGH NE"

BASE COUNT 519 a 735 c 689 g 634 t

ORIGIN

Query Match 56.9%; Score 1454.8; DB 10; Length 2577;

Best Local Similarity 74.3%; Pred. No. 3.2e-185;

Matches 1891; Conservative 0; Mismatches 622; Indels 33; Gaps 3;

QY 1 ATGCTGGGCCCTGCTGCTGGGCCCTGAGCCTGCTCTGCAACCTGGAGCGGG 60

DB 1 ATGCGACCTTTGGCTATCATGAGGTCTGAGCCTGCTCTTCTGGAGCTGGATGGG 60

QY 61 GCCCATTTGGCTGTACAGCACTTACGATGAGGGGACTACGCTGGGGGGCTG 120

DB 61 GCCCTTTTGGCTGTACAGCACTTACGATGAGGGGACTACGATCTGGGGGGCTA 120

QY 121 TTCCGCTGGGGGAGGAGGAGGCTGGGCTCCGAGCGCGGACAGCGGAGCCT 180

DB 121 TTTCCTCTGGGCTACCGAGGAGGAGGCTCTCAACGAGAGACCAACCAAGCATC 180

QY 181 GTGTGACACAGTTCTCTCAAAAGCGCTGCTGTGGACATGGCCATGAATATGCCGTG 240

DB 181 CGGTGCAACAGGTCTTCACCCCTGGTTTGTCTGTGGCCATGSCATFMAAGATGGCTG 240

QY 241 GAGGAGATCAACAACAAGTCGATCTGCTGCCGGGCTGCGCTGGGCTACACCTTT 300

DB 241 GAGGAGATCAACAACAAGTGAATGCTGCTGCCGGGCTGCGGCTGGGCTATGACTATTT 300

QY 301 GATACGGTCGCGGAGGCTGTGGGCCATGAGACCCAGCCTCATGTTCCGGGCAAGGCA 360

DB 301 GACACATGCTCCGAGCGCAGTGTACACCAATGAATTCAGTCTCATGTTCCGGGCAAGG 360

QY 361 GGCAGCGGCACATCGCCGCTACTGCACTACACGACGATGACAGCCCGCTGTGCTGCT 420

DB 361 GGCAGTCACAGCATTTGTGCTACTGCACTACACAGATGACACAGACCCGCTGTGCTGCT 420

QY 421 GTCATGGGGCCCACTGCTGACAGCTGCGCATGTGTCACCGCAAGTTCTTCACTTCTTC 480

DB 421 GTCATGGGGCCCACTGCTGACAGCTGCGCATGTGTCACAGCAAGTTCTTCACTTCTTC 480

QY 481 CTCATGCCCCAGGTCAGCTACGCTGCTAGCATGAGCTGTGAGCGCCGAGACCTTC 540

DB 481 CTCATGCCCCAGGTCAGCTACGCTGCTAGCATGAGCTGTGAGCGCCGAGAAAGTTT 540

QY 541 CCCTCTCTTTCGCAACCGTGCACGACGAGCTGTGCACTGACGTGACGCCCGGAGCTG 600

DB 541 CCATCTCTTTCGCAACAGGTCGACGAGCAGCGGCTGAGCTGACGAGCAAGTTGTACTCTG 600

QY 601 CTCGAGAGTTGGGCTGGAAGTGGTGGCGGCTGGGGGCGGACGACGAGTGGGCGG 660

DB 601 TTGCGAAGCTTCACTGGAAGTGGTGGCGGCTGGGGGCGGACGAGTGGGCGG 660

QY 661 CAGGCGCTGACACATCTCTGCGGCTGGCGGCGGACGCGGCACTGTCATCGGCGAG 720

DB 661 GAGGCTGACACATCTTCTTCTGCTGCGGCAATGACAGAGTATCTGCAATCGGCAAT 720

QY 721 GGCTGTGGCGCTGCCCCGCTGCGGATGACTCGGCGCTGGGGAAGGTGACAGCTCTG 780

DB 721 GGCTGTGGCGCAACATGACACTAGTGGGCAACAGTGGGCAAGGTGCTGAGTACTA 780

QY 781 CACCAAGTGAACACAGACAGCTGACAGTGTGCTGTGCTGCTGCGGCGGCGGCG 840

DB 781 CGGCAAGTGAACACAAATGAAGTGAAGTGTGCTGCTGCTGCGGCGGCGGCG 840

QY 841 CAGCCCTCTTCAACATGACATGACGACGAGCGCTGCGGCAAGGTGGTGGCGAGC 900

DB 841 TACTCCCTTTTACTTACAGATCATCATGAGCTCTACACCAAGATATGGTGGCGAGT 900

QY 901 GAGGCTGTGCTGACCTGCTGACCTGTGATGAGGGCTGCGGCGGACATGCGCAATGGCGAG 960

DB 901 GAGCTGTGCTGACATGCTGACCTGTGATGACCTGCAATATGGCGGCTGTGGCGAGT 960

QY 961 GTGTTGGCTTCCCGCAGAGGGGTGCCACGTGACAGAGTTCGCCCATGCGGGAAGC 1020

DB 961 GTGTTGGGTTTTTTCAGCGGGGTGCCCTACTGCTCAATATTTTCCCAATATATGGAGACT 1020

QY 1021 CACCTGCGCTGCGACCGACCGGCGCTTGTGCTGCTGCGGCGGAGAGGAGCAGAGT 1080

DB 1021 CACCTTCCCTGCGCGGTGACCGACGATCTGTGCTGCTGCTGCGGAGAGTGGAT 1077

QY 1081 CTGAGGAGAGAGCTGTGGGCGCAGCGCTGCGGAGTGTACATGATCAGC----- 1132

DB 1078 CTGAGGAGAGAGTGTGGGCGCAGCGCTGCTGCAAGGTGTACGACATGATGCTGACAGAC 1137

QY 1133 -----TGCAGAACGTGAGGCGAGGGCTTAATGATCAACGACGAGCTTCT 1176

DB 1138 CTATATCTGTGGCTGTGTGCAAGAACTATGCACTGTGGGAAATTTGCAACCAAAATTTTGA 1197

QY 1177 GTCTACGACAGCTGTATGATGCTGTGGCCAGGCGCTGCAACAACCTTCTCAAGTGAAGCC 1236

DB 1198 ACCTATGACAGCTGTATGATGCTGTGGCTCAAGGCTTTCACAACAACCTTCAAGTGAAGTCT 1257


```

OY 1237 TCAGGCTGCCCGCCAGAGACCCCTGGAAGCCCTGCGACCTCTGGAGAACATGTACAAC 1296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1238 TCACATTCGACAGATATACAGAACATTTCTACCTTGCGACCTCTCTGGAGAACATGTACAAT 1317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1237 CTGACCTTCCACGTGGGGGGGGCTGCGGCTTCGACAGCAGGGAACGTGACATG 1356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1318 ATGATTTCCATGCTCGAGACTTGTACACTACAGTTTGTATGTGAAGGAATGTACATG 1377
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1357 GAGTACGACCTGAGCTGGGGTGTGGCAGGGCTCAGTGCCCAAGCTCCACAGCTGGGG 1416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1378 GATATGTACCTGAAATGTGGGTGTGGCAGAGCCCTACACCTGTATTTACATCTGTGGC 1437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1417 AGGTTCAAGCGCAGCCTCAGACAGAGCCCTGAAATCCGTGGCAGACGCTACACAAC 1476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1438 ACCTTCAGAGGACCCCTTACGCTGACGACGCTGTAAAATGTACTGGC-----CAGGCAAC 1491
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1477 CAGAAAGCCGCTGCCGCTGCTCGCGGAGTCCAGAGAGGGCCAGGTGCGCCGGCTCAAG 1536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1492 CAGGTGCGCAGTCTCCAGAGTTCGCCGACAGTCAAAAGATGGCCAGGTTCCGCGAGTAAAG 1551
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1537 GGGTTCACACTCTGCTGCTACAGACTGTGTGACCTCGAGAGGGGCGACGTACCGGCAAAAC 1596
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1552 GGGTTTCATCTCTGTGTATGACTGTGCTGCAAGGCGGCGACGCTACCGGAAGCAT 1611
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1597 CCAGAGCAGATCGCCTGACACTTTTGTGGCCAGATGATGATGCTCCCGGAGGCAAGCACA 1656
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1612 CCAGATGATCTACCTGTACTGTATCAATGTAACAGAGACAGTGTCCCGCAGAGAAAGCACA 1671
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1657 CGCTGCTTCCGCGCAGGTCTCGGTTCTGSGCATGGGGGAGCCGAGTGTGCTGCTG 1716
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1672 GCTGCTTACCTCGCAGAGGCCCAAGTTTCGCTGGGGGAGCCAGTGTGCTGTCACTC 1731
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1717 CTGCTGCTGCTAGGCTGGGCGCTGTGCTGGGCTGTGGGCTGTGGGCTGTGCTGCTG 1776
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1732 CTGCTGCTGCTTGGCTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1777 CATCGGAGACGCCCACTGTTTCAGGCTCGGGGGGGCCCTGCGCTGTTGGCTGTG 1836
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1792 CACTGGAGACGCCCTGTGTTCAGAGCCTCAGGTGAGTGTGAGTGTGAGTGTGAGTGTG 1851
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1837 TGCTTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1896
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1852 TGCTTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1911
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1897 TGCTTGGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1956
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1912 TGCTTGGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1971
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1957 CTGACAGGCGCCGAGATCTTCTGAGTCAAGATGCTGCTGAGCTGGGAGAGCCGGCTG 2016
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1972 CTGACAGGCGAGTGAACCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2017 AGTGGCTGCTGCGGGGGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 2076
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2032 TGCACTTACCTTCTGGGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 2091
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2077 GTGCACTGTGACCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2092 GAGGACATGTGCTGTGATTTGATTCGCTTTCACACAGAGTGTGTGACAGACTGTGCA 2151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2137 ATGCTGCCACGAGGCGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2152 GTGCTGCCACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2197 GCGCAGCGCAGATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2212 GTGCACTGTACCAATGTGATGTGATGTGATGTGATGTGATGTGATGTGATGTGATG 2271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2257 AGCCAGCGGCGCTGTGACAAAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2272 AGCCAGCGGCGCTGTGACAAAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2317 ATCAGCTGGGTCTCTTGTGCGCTCTGCGCAATGTGACAGTGTCTCTGAGGCCCGCC 2376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 2332 ATCAGCTGGGTCTCTTGTGCGCTCTGCGCAATGTGACAGTGTCTCTGAGGCCAGT 2391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2377 GTGCAATGTGCGGCCCTCTCTGTGTGCTGTGAGCATCTGTGCTTCCACCTGCCC 2436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2392 GTGCAATGTGCGGCCCTCTCTGTGTGCTGTGAGCATCTGTGCTTCCACCTGCCC 2451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2437 AGTGTACCTGCTCATGCGGAGCGCAACACCCGAGTGTCTCTCTGAGGAGG 2496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2452 AAGTGTATGTCTCTTGTGCGCAAGCTCAACACCCAGAGTGTCTCTCTGAGGAGG 2511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 2497 GGGCTTGGGATGCCCAAGGCCAGAA 2522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2512 AATGCCAAGAAAGCAGAGATGAGAA 2537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: May 23, 2003, 16:46:17
Job time : 7172.01 secs

THIS PAGE BLANK (USPTO)

QY 1561 TGTGTGACACTGAGAGCGGAGAGCTACCGGAAAAACCGAGACATGCTGACCTTT 1620
 DB 1561 TGTGTGACACTGAGAGCGGAGAGCTACCGGAAAAACCGAGACATGCTGACCTTT 1620
 QY 1621 TGTGTGACACTGAGAGCTGCTCCCGAGAGCAACAGCTGCTCCCGAGAGCTG 1680
 DB 1621 TGTGTGACACTGAGAGCTGCTCCCGAGAGCAACAGCTGCTCCCGAGAGCTG 1680
 QY 1681 TGTGTGACACTGAGAGCTGCTCCCGAGAGCAACAGCTGCTCCCGAGAGCTG 1740
 DB 1681 TGTGTGACACTGAGAGCTGCTCCCGAGAGCAACAGCTGCTCCCGAGAGCTG 1740
 QY 1741 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
 DB 1741 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
 QY 1801 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
 DB 1801 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
 QY 1861 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
 DB 1861 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
 QY 1921 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
 DB 1921 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
 QY 1981 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
 DB 1981 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
 QY 2041 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
 DB 2041 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
 QY 2101 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
 DB 2101 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
 QY 2161 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
 DB 2161 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
 QY 2221 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
 DB 2221 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
 QY 2281 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
 DB 2281 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
 QY 2341 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
 DB 2341 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
 QY 2401 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
 DB 2401 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
 QY 2461 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
 DB 2461 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
 QY 2521 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
 DB 2521 GGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580

RESULT 2
 AAL38462
 ID AAL38462 standard; cdna: 3458 BP.
 XX

AC AAL38462;
 XX 15-AUG-2002 (first entry)
 DE cDNA encoding a human G-protein coupled receptor (GPCR).
 DE G-protein coupled receptor; GPCR; gene chip; human; immune response;
 KW chromosomal position; transgenic animal; gene therapy; gene; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 30..2588
 FT /tag= a
 FT /product= "Human G-protein coupled receptor"
 XX
 XX WO20030981-A1.
 XX 18-APR-2002.
 XX 13-MAR-2001; 2001WO-US07832.
 XX 10-OCT-2000; 2000US-0684393.
 XX (PEKE) PE CORP NY.
 XX Wei M, Zhong W, Ketchum KA, Difrancesco V, Beasley EM;
 DR WPI: 2002-444173/47.
 DR P-PSDB; AAO21501.
 PT Novel G protein coupled receptor, useful for raising antibodies, to
 PT elicit immune response, and as a reagent in assays designed to
 PT quantitatively determine levels of protein in biological samples -
 XX
 XX Claim 1; Fig 1; 82pp; English.
 XX
 XX The invention relates to an isolated G-protein coupled receptor (GPCR)
 XX polypeptide, comprising an 852 residue amino acid sequence, given in the
 XX specification, an allelic variant or ortholog of the protein, or a
 XX fragment comprising at least 10 contiguous amino acids of the protein.
 XX GPCR is useful for identifying a modulator of GPCR and an agent that
 XX binds to GPCR. GPCR and a gene chip comprising GPCR are useful as models
 XX for the development of human therapeutic agents. GPCR is useful for
 XX raising antibodies, to elicit immune response, as a reagent in assays
 XX designed to quantitatively determine levels of protein in biological
 XX samples, and as markers for tissues in which the corresponding protein
 XX is preferentially expressed. A gene chip containing GPCR is also useful
 XX as a probe for determining the chromosomal positions of nucleic acid
 XX molecules by means of in situ hybridisation, in making vectors containing
 XX the gene regulatory regions of a gene chip containing GPCR, for designing
 XX rhizomes, in making vectors that express GPCR, and for constructing host
 XX cells and transgenic animals expressing nucleic acid molecules and
 XX peptides. A host cell containing GPCR is useful for conducting cell-based
 XX assays involving GPCR protein or its fragments, and for identifying
 XX protein mutants in which these function is affected. The polynucleotide
 XX encoding GPCR can be used to treat disorders by gene therapy. This
 XX polynucleotide sequence represents cDNA encoding the human G-protein
 XX coupled receptor of the invention.
 XX
 XX Sequence 3458 BP; 590 A; 1179 C; 1092 G; 597 T; 0 other;
 S0
 Query Match 99.9%; Score 2557.4; DB 24; Length 3458;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2558; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGCTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
 DB 30 ATGCTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 89
 QY 61 GCCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 DB 90 GCCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 149

OY	121	TTCCCCCTGGGGCAGGGCCGAGGAGGGGCTGGCTCCGAGGCGGACAGGGCCAGAGCCCT	180
OY	121	TTCCCCCTGGGGCAGGGCCGAGGAGGGGCTGGCTCCGAGGCGGACAGGGCCAGAGCCCT	180
Db	150	TTCCCCCTGGGGCAGGGCCGAGGAGGGCTGGCTCCGAGGCGGACAGGGCCAGAGCCCT	209
OY	181	GTGTGCACACAGTTCTCTCTCAAAAGGGCTGCTGTGGGCACTGGACATGAAATGGCCGTG	240
Db	210	GTGTGCACACAGTTCTCTCTCAAAAGGGCTGCTGTGGGCACTGGACATGAAATGGCCGTG	269
OY	241	GAGGAGATCAACAACAGTGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	300
Db	270	GAGGAGATCAACAACAGTGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	329
OY	301	GATACGTCTCGGAGCCCTGT	360
Db	330	GATACGTCTCGGAGCCCTGT	389
OY	361	GGCAGCCGCGACATCGCCGCCCTACTGTCACTACAGCAGTACCAAGCCCGTGTGCTGCT	420
Db	390	GGCAGCCGCGACATCGCCGCCCTACTGTCACTACAGCAGTACCAAGCCCGTGTGCTGCT	449
OY	421	GTCAATCGGGCCCCCTGT	480
Db	450	GTCAATCGGGCCCCCTGT	509
OY	481	CTCATGCCCCCAGGTCACTACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	540
Db	510	CTCATGCCCCCAGGTCACTACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	569
OY	541	CCCTCTCTTCTTCGCGACCCGTCGCGACGCGTGTGTGTGTGTGTGTGTGTGTGTGTGT	600
Db	570	CCCTCTCTTCTTCGCGACCCGTCGCGACGCGTGTGTGTGTGTGTGTGTGTGTGTGTGT	629
OY	601	CTGACAGAGTTCCGCTGTGAATCGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	660
Db	630	CTGACAGAGTTCCGCTGTGAATCGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	689
OY	661	CAGGCGCTGAGCATCTTCTGTGGCCCTGTGGCCGCGGACAGCGGCACTGTGCATGTGCG	720
Db	690	CAGGCGCTGAGCATCTTCTGTGGCCCTGTGGCCGCGGACAGCGGCACTGTGCATGTGCG	749
OY	721	GACCGGGGCGCGTGGCCCGGCGGCGATGTGCGGGCTGGGGAAAGGTGTGAGGACGTCCT	780
Db	750	GACCGGGGCGCGTGGCCCGGCGGCGATGTGCGGGCTGGGGAAAGGTGTGAGGACGTCCT	809
OY	781	CACCAAGTGAAACAGAGCAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	840
Db	810	CACCAAGTGAAACAGAGCAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	869
OY	841	CACGCGCTCTTCAACTACAGCATCAGCAGCAGGCTTCGCGCAAGGTGTGGTGGCCAGC	900
Db	870	CACGCGCTCTTCAACTACAGCATCAGCAGCAGGCTTCGCGCAAGGTGTGGTGGCCAGC	929
OY	901	GAGGCGTGGGCGACCTGTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	960
Db	930	GAGGCGTGGGCGACCTGT	989
OY	961	GTGCTTGGCTTCTCCAGAGGGGTGCCACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1020
Db	990	GTGCTTGGCTTCTCCAGAGGGGTGCCACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1049
OY	1021	CACCTGGCCCTGGGCACCGACCCGGCCCTTGTGCTGTGCCCTGTGGCGAGAGGAGCAGGG	1080
Db	1050	CACCTGGCCCTGGGCACCGACCCGGCCCTTGTGCTGTGCCCTGTGGCGAGAGGAGCAGGG	1109
OY	1081	CTGAGAGAGAGCATGTGTGTGGGCGACGCGTGTGCGCACTGTGACTCATCAGCGTGTGCA	1140
Db	1110	CTGAGAGAGAGCATGTGTGTGGGCGACGCGCGCTGTGCGCACTGTGACTCATCAGCGTGT	1169
OY	1141	GTGAGCCAGAGGCTTAATTCACCAACAGACCTTCTGTGTGTGTGTGTGTGTGTGTGTGT	1200
Db	1170	GTGAGCCAGAGGCTTAATTCACCAACAGACCTTCTGTGTGTGTGTGTGTGTGTGTGTGT	1229

QY	1201	GCACGAGCCCTGCACAACTCTTTCAGTGCACAGCCCTCAGGGCTGCCCGCGAGAGACCCC	1260
Db	1230	GCCACGAGCCCTGCACAACTCTTTCAGTGCACAGCCCTCAGGGCTGCCCGCGAGAGACCCC	1289
QY	1261	GTGAAGCCCTTGCGACGCTCTGAGAAACATGTACAACTTGACCTTCCACGTGAGCGGGCTG	1320
Db	1290	GTGAAGCCCTTGCGACGCTCTGAGAAACATGTACAACTTGACCTTCCACGTGAGCGGGCTG	1349
QY	1321	CCGCTGGGGGTTCCGACACACACGGGAAACGTGGGACATGAGATACGACTGAGCTGGGGTG	1380
Db	1350	CCGCTGGGGGTTCCGACACACACGGGAAACGTGGGACATGAGATACGACTGAGCTGGGGTG	1409
QY	1381	TGGCAGGAGCTCAGTGGCCAGGCTCCACGACGTGGGAGGTTCCAMGCGAGCCTCAGGACA	1440
Db	1410	TGGCAGGAGCTCAGTGGCCAGGCTCCACGACGTGGGAGGTTCCAMGCGAGCCTCAGGACA	1469
QY	1441	GAGGCGCTGAAAGATCCGCTGGGACACGTGTGCACAAACGAAACCCGTTGCCGGTGTGCG	1500
Db	1470	GAGGCGCTGAAAGATCCGCTGGGACACGTGTGCACAAACGAAACCCGTTGCCGGTGTGCG	1529
QY	1501	CGCGAGTGCAGAGAGGGGCCAGGTGCCCGGGGTCACATCCTGCTGCTAGAC	1560
Db	1530	CGCGAGTGCAGAGAGGGGCCAGGTGCCCGGGGTCACATCCTGCTGCTAGAC	1589
QY	1561	TGTGTGACACTGCGAGGCGGGCACCTTACCGGAAACCAGACGACATGCTGCTGACCTTT	1620
Db	1590	TGTGTGACACTGCGAGGCGGGCACCTTACCGGAAACCAGACGACATGCTGCTGACCTTT	1649
QY	1621	TGTGGCCAGGATGTAGTGTGTCCCGGAGACGAACACGCTGTTCCGGCCGAGGCTTCG	1680
Db	1650	TGTGGCCAGGATGTAGTGTGTCCCGGAGACGAACACGCTGTTCCGGCCGAGGCTTCG	1709
QY	1681	TTCTGCGCATGAGGGCGAGCCGGCTGTGCTGTCTGCTCTGCTGCTGAGCCTGAGCGCTG	1740
Db	1710	TTCTGCGCATGAGGGCGAGCCGGCTGTGCTGTCTGCTCTGCTGCTGAGCCTGAGCGCTG	1769
QY	1741	GGCCTTGAGCGAGGCTGATTTGGGGGCTTGCTTCAACCAATCGGGACAGCCACTGAGTTACG	1800
Db	1770	GGCCTTGAGCGAGGCTGATTTGGGGGCTTGCTTCAACCAATCGGGACAGCCACTGAGTTACG	1829
QY	1801	GCCTCGAGGGGGGCCCTGAGCCTCTTTGGGCTGTGTGCTGCTGGGCTGTGTGCTTCAGC	1860
Db	1830	GCCTCGAGGGGGGCCCTGAGCCTCTTTGGGCTGTGTGCTGCTGGGCTGTGTGCTTCAGC	1889
QY	1861	GTCCTCCGTTTCCCTGGGACAGCCACACCCCGCCATGCTGTGGCCACACAGCCCTGTGTC	1920
Db	1890	GTCCTCCGTTTCCCTGGGACAGCCACACCCCGCCATGCTGTGGCCACACAGCCCTGTGTC	1949
QY	1921	CACCTCCCGGTCACAGGGGCTCCCTGAGCACACTTTCCTGACAGGCGGGCAGAGATCTTGCTG	1980
Db	1950	CACCTCCCGGTCACAGGGGCTCCCTGAGCACACTTTCCTGACAGGCGGGCAGAGATCTTGCTG	2009
QY	1981	GAGTCACAAACTGCTCTGACACTTGGGACAGACCGGCTGTGAGTGCTGCTCGGGGGCCCTG	2040
Db	2010	GAGTCACAAACTGCTCTGACACTTGGGACAGACCGGCTGTGAGTGCTGCTCGGGGGCCCTG	2069
QY	2041	GCCGCGGTGGGGGTCGTCGGGCAATGCGTGGGAGAGTGCACATGTGACACTGGGACTCTG	2100
Db	2070	GCCGCGGTGGGGGTCGTCGGGCAATGCGTGGGAGAGTGCACATGTGACACTGGGACTCTG	2129
QY	2101	GTGAGCTTTCGCCGAGAGGTGTGAGTGGAGTGGCAATGCTCCACACGAGAGGCTGCTGGTG	2160
Db	2130	GTGAGCTTTCGCCGAGAGGTGTGAGTGGAGTGGCAATGCTCCACACGAGAGGCTGCTGGTG	2189
QY	2161	CAGTGCACACAGCTTCCTGAGGCTTGCGCTTACGCGACGCCCAATGCCACGCTG	2220
Db	2190	CAGTGCACACAGCTTCCTGAGGCTTGCGCTTACGCGACGCCCAATGCCACGCTG	2249
QY	2221	GCCCTTTCGTCGCTCCGAGGCAATTTTCGTTGGGCGAGAGCAGCCGGGCTGTACAAACGCT	2280
Db	2250	GCCCTTTCGTCGCTCCGAGGCAATTTTCGTTGGGCGAGAGCAGCCGGGCGCTACAAACGCT	2309
QY	2281	GCCCGTGGCCCTACCTTGTGCATGCTGGCTACTTATCACCCTGGGTCCTCTTGTGACC	2340

Db 2310 GCGCGTGGCTCACCCTTGGCTGCTGCTACTCATCACCCTGGGTCTCCCTTGTGCC 2369
 Qy 2341 CTCCTGGCCATGTCAGAGTGGTCTCTCAGAGCCCGCTGCAGATGGCCCTCTGCTC 2400
 Db 2370 CTCCTGGCCAAATGTCAGAGTGGTCTCTCAGAGCCCGCTGCAGATGGCCCTCTGCTC 2429
 Qy 2401 TGTGCTGGGCTGCTGCTGCTCTTCCACGCGCCAGAGTGTACTGCTCATGGGCGAG 2460
 Db 2430 TGTGCTGGGCTGCTGCTGCTCTTCCACGCGCCAGAGTGTACTGCTCATGGGCGAG 2489
 Qy 2461 CCAGGGCTCAACACCCCGAGTTCTTCTGGAGGGGCGCTGGGAGATGCCCAAGGCCAG 2520
 Db 2490 CCAGGGCTCAACACCCCGAGTTCTTCTGGAGGGGCGCTGGGAGATGCCCAAGGCCAG 2549
 Qy 2521 AATGACGGGAAACAGAAATCAGGGGAAACATGAGTGA 2559
 Db 2550 AATGACGGGAAACAGAAATCAGGGGAAACATGAGTGA 2588

RESULT 3

AAS97395 standard; cDNA: 2559 BP.

AAS97395;

12-MAR-2002 (first entry)

Human SACL gene cDNA.

Human; mouse; SACL; carbohydrate; sweetener; ethanol; alcoholism; ss;

obesity; diabetes; transgenic embryo; body tissue; body fluid; pancreas;

blood; tongue; PCR primer; anorectic; antidiabetic; gene therapy;

protein replacement therapy.

Homo sapiens.

WO200183749-A2.

08-NOV-2001.

25-APR-2001; 2001WO-US13387.

28-APR-2000; 2000US-200794P.

28-JUL-2000; 2000US-221419P.

10-NOV-2000; 2000US-247443P.

(WARN) WARNER LAMBERT CO.

(MONE-) MONELL CHEM SENSES CENT.

Bachmanov AA, Beauchamp GK, Chatterjee A, De Jong PJ, Li S, Li X;

Ohmen JD, Reed DR, Ross D, Tordoff MG;

WPI, 2002-075162/10.

P-PSDB; AAU73184.

Novel isolated polypeptide comprising variant form of mouse or human

SACL polypeptide, and is associated with altered preference for

carbohydrates or other sweeteners, useful for preventing obesity,

diabetes, alcoholism

Claim 3; Page 42-44; 239pp; English.

CC (such as blood, pancreas or tongue) for sequence variations of the SACL
 CC gene. A sequence variation of the SACL locus may indicate a
 CC predisposition to diabetes, obesity and/or alcoholism and may provide a
 CC diagnostic mark. The polynucleotide can be detected in a biological
 CC sample by contacting the DNA with a probe to form a hybridisation complex
 CC which is then detected. The sequences represent cDNA encoding human and
 CC mouse SACL polypeptides and PCR primers specific for the SACL genes.
 XX

Sequence 2559 BP; 380 A; 863 C; 840 G; 476 T; 0 other;

Query Match 99.9%; Score 2555.8; DB 24; Length 2559;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGCTGGGCGCCCTGCTGCTCTGAGCTCAGCCCTCTGCTCTCTCAGCCTTGGAGCGGG 60
 Db 1 ATGCTGGGCGCCCTGCTGCTCTGAGCTCAGCCCTCTGCTCTCTCAGCCTTGGAGCGGG 60
 Qy 61 GCGCCATTGTGCTCTCAGCACTTAGATGAAGGAGAACTAGTCTGGAGGGGCTG 120
 Db 61 GCGCCATTGTGCTCTCAGCACTTAGATGAAGGAGAACTAGTCTGGAGGGGCTG 120
 Qy 121 TTCCCTGGGCGAGGCGGAGGAGGCTGCTCCGACCGGACAGGCGCCAGACCCCT 180
 Db 121 TTCCCTGGGCGAGGCGGAGGAGGCTGCTCCGACCGGACAGGCGCCAGACCCCT 180
 Qy 181 GTGTGACACAGGTCTCTCTCAAAAGGCGCTGCTCTGGGCACTGAGGCAATGAAATGGCCGTG 240
 Db 181 GTGTGACACAGGTCTCTCTCAAAAGGCGCTGCTCTGGGCACTGAGGCAATGAAATGGCCGTG 240
 Qy 241 GAGGAGATCAACAACAAGTGGATCTGTCGCCGGGCTGCGCTGGGCTAGACACTTTT 300
 Db 241 GAGGAGATCAACAACAAGTGGATCTGTCGCCGGGCTGCGCTGGGCTAGACACTTTT 300
 Qy 301 GATACGTCTCTGGAGCTGTGTGGCCATGAAGCCAGCTTCAATGTTCTTGCCAGAGCA 360
 Db 301 GATACGTCTCTGGAGCTGTGTGGCCATGAAGCCAGCTTCAATGTTCTTGCCAGAGCA 360
 Qy 361 GGCACCGCGGACATGCGCGGCTAGTCACTACAGAGTACAGGCGCGGCTGTGGCT 420
 Db 361 GGCACCGCGGACATGCGCGGCTAGTCACTACAGAGTACAGGCGCGGCTGTGGCT 420
 Qy 421 GTCATGGGCGCCCACTGCTCAGAGCTGCCATGATGTCACCGGCAATGTTCTTACGTTCTTC 480
 Db 421 GTCATGGGCGCCCACTGCTCAGAGCTGCCATGATGTCACCGGCAATGTTCTTACGTTCTTC 480
 Qy 481 CTCATGCCCCAGGTACGTAGGCTCTAGCATGAGCTGTGAGCGCCCGGAGACCTTTC 540
 Db 481 CTCATGCCCCAGGTACGTAGGCTCTAGCATGAGCTGTGAGCGCCCGGAGACCTTTC 540
 Qy 541 CCTCTCTTCCGACGCTGCGCCAGAGCGTGTGAGCTGAGCGCGCGCGGAGAGCTG 600
 Db 541 CCTCTCTTCCGACGCTGCGCCAGAGCGTGTGAGCTGAGCGCGCGCGGAGAGCTG 600
 Qy 601 CTGCAGAGTTGCGCTGGAACCTGGGTGGCCCGCTTGGGCGAGGAGATACGCGCGG 660
 Db 601 CTGCAGAGTTGCGCTGGAACCTGGGTGGCCCGCTTGGGCGAGGAGATACGCGCGG 660
 Qy 661 CAGGCGCTGAGCATTTCTTGGCGCTGCGCGGCGGACGCGGCAATGATGCGGACAG 720
 Db 661 CAGGCGCTGAGCATTTCTTGGCGCTGCGCGGCGGACGCGGCAATGATGCGGACAG 720
 Qy 721 GCGCTGGTGGCGCGCGCGCGCGCGCATACGCGGCGGGAAGTCAAGACGCTCTG 780
 Db 721 GCGCTGGTGGCGCGCGCGCGCGCGCATACGCGGCGGGAAGTCAAGACGCTCTG 780
 Qy 781 CACCAAGTGAACAAGAGAGGCTGAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 Db 781 CACCAAGTGAACAAGAGAGGCTGAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 Qy 841 CACGCGCTCTTCACTACAGCATGACAGAGGCTCTGCGGCAAGTGTGGTGGCGAGC 900
 Db 841 CACGCGCTCTTCACTACAGCATGACAGAGGCTCTGCGGCAAGTGTGGTGGCGAGC 900

PN WO200164882-A2.
XX 07-SEP-2001.
XX 28-FEB-2001; 2001WO-US06543.
XX 29-FEB-2000; 2000US-186059P.
XX (MILL-) MILLENNIUM PHARM INC.
XX Glucksmann MA, Galvin KM, Silos-Santiago I;
XX WPI: 2001-589866/66.
XX P-PSDB: MAU08996.
PT Novel G protein coupled receptors and nucleic acids encoding them, for
PT identifying agents for the treatment of cardiac disorders -
PS Claim 2; Fig 15; 209pp; English.
XX The invention relates to novel human G protein-coupled receptors (GPCR)
CC named 1983, 52881, 2398, 45449, 50289 and 52872. The novel GPCRs and
CC nucleic acids encoding them are useful for identifying agents for the
CC treatment of cardiovascular disorders, angiogenesis-related disorders,
CC neutral disorders, pain response disorders and inflammatory disorders
CC e.g. atherosclerosis, angina pectoris and myocardial infarction,
CC ischaemic heart disease, sudden cardiac death, hypertensive heart
CC disease, diabetes, prostate cancer-related pain, diabetes and obesity.
CC The present sequence encodes GPCR 50289.
XX
XX Sequence 3489 BP; 578 A; 1189 C; 1102 G; 612 T; 8 other;
Query Match 99.9%; Score 2555.8; DB 22; Length 3489;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGTGGGCGCTGTCTGCTGGGCTCAAGCTTGGGCTCTCTGCTGACCTGGAGCGGG 60
DB 52 ATGTGGGCGCTGTCTGCTGGGCTCAAGCTTGGGCTCTCTGCTGACCTGGAGCGGG 111
QY 61 GCGCCATTGTGCTGCTCAAGCACTTAGTAGAGGGGAGTACTGTGGGGGGGCTG 120
DB 112 GCGCCATTGTGCTGCTCAAGCACTTAGTAGAGGGGAGTACTGTGGGGGGGCTG 171
QY 121 TTCCCTGAGGAGGAGGAGGAGGCTGCTCCGACGCGGACAGGCGGACAGCTT 180
DB 172 TTCCCTGAGGAGGAGGAGGAGGCTGCTCCGACGCGGACAGGCGGACAGCTT 231
QY 181 GTGTGCACCAAGTCTCTCAAAAGGCTGCTGTGGGCACTGGCATGAATAATGGCCGTG 240
DB 232 GTGTGCACCAAGTCTCTCAAAAGGCTGCTGTGGGCACTGGCATGAATAATGGCCGTG 291
QY 241 GAGGAGATCAACAAGCGGATGCTGCTCCCGGGCTGGCGCTGAGTACAGCTCTTT 300
DB 292 GAGGAGATCAACAAGCGGATGCTGCTCCCGGGCTGGCGCTGAGTACAGCTCTTT 351
QY 301 GATACGTGCTGAGGCTGTGTGGGCAATGAAGCCAGGCTCATTTCTGAGCAAGCA 360
DB 352 GATACGTGCTGAGGCTGTGTGGGCAATGAAGCCAGGCTCATTTCTGAGCAAGCA 411
QY 361 GCGACCGCGGAGATGCGCGCTACTGTCAATACAGCAGATACAGCCCGCTGTGCTG 420
DB 412 GCGACCGCGGAGATGCGCGCTACTGTCAATACAGCAGATACAGCCCGCTGTGCTG 471
QY 421 GTGATCGGGGCGGCTGCTGAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 472 GTGATCGGGGCGGCTGCTGAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 531
QY 481 CTGATCGGGGCGGCTGCTGAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 532 CTGATCGGGGCGGCTGCTGAGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 591
QY 541 CCTCTCTTCTTCCGACCGCTGCCAGCAGCGTGTGCTGCTGCTGCTGCTGCTGCTG 600

DB 592 CCTCTCTTCTTCCGACCGCTGCCAGCGACCGCTGTGCTGCTGCTGCTGCTGCTG 651
QY 601 CTGACAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 652 CTGACAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 711
QY 661 CAGGCGCTGAGCATCTTCTGCGGCTGCGGCGGAGCAGCGGATGCTGCTGCTGCTG 720
DB 712 CAGGCGCTGAGCATCTTCTGCGGCTGCGGCGGAGCAGCGGATGCTGCTGCTGCTG 771
QY 721 GCGCTGTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 780
DB 772 GCGCTGTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 831
QY 781 CACGAGGTGAACCAAGAGCAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 832 CACGAGGTGAACCAAGAGCAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 891
QY 841 CAGGCGCTTCTTCAACTACAGCATACAGCAGGCTGCTGCGGAGTGTGGTGGCAGC 900
DB 892 CAGGCGCTTCTTCAACTACAGCATACAGCAGGCTGCTGCGGAGTGTGGTGGCAGC 951
QY 901 GAGGCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB 952 GAGGCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1011
QY 961 GTGCTGTGCTTCTTCAAGAGGCTGCGGAGCAGGCTGCTGCTGCTGCTGCTGCTG 1020
DB 1012 GTGCTGTGCTTCTTCAAGAGGCTGCGGAGCAGGCTGCTGCTGCTGCTGCTGCTG 1071
QY 1021 CACCTGTGCTTCTTCAAGAGGCTGCGGAGCAGGCTGCTGCTGCTGCTGCTGCTG 1080
DB 1072 CACCTGTGCTTCTTCAAGAGGCTGCGGAGCAGGCTGCTGCTGCTGCTGCTGCTG 1131
QY 1081 CTGAGAGGAGAGCTGTGTGGGCGGAGCGCTGCGGAGTGTGCTGCTGCTGCTG 1140
DB 1132 CTGAGAGGAGAGCTGTGTGGGCGGAGCGCTGCGGAGTGTGCTGCTGCTGCTG 1191
QY 1141 GTGAGGAGGAGGCTTAAATACACAGCAGGCTGCTGCTGCTGCTGCTGCTGCTG 1200
DB 1192 GTGAGGAGGAGGCTTAAATACACAGCAGGCTGCTGCTGCTGCTGCTGCTGCTG 1251
QY 1201 GCCAGGCGCTTCAACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
DB 1252 GCCAGGCGCTTCAACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1311
QY 1261 GTGAGGCGCTTCAACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB 1312 GTGAGGCGCTTCAACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1371
QY 1321 CGGCTGTGCTTCAACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
DB 1372 CGGCTGTGCTTCAACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1431
QY 1381 TGGCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB 1432 TGGCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1491
QY 1441 GAGGCGCTTCAACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
DB 1492 GAGGCGCTTCAACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1551
QY 1501 GCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
DB 1552 GCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1611
QY 1561 TGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
DB 1612 TGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1671
QY 1621 TGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680

[illegible]

RESULT 5	
AAD17509	
ID	AAD17509 standard; cDNA; 2553 BP.
XX	
AC	
XX	AAD17509;
DT	10-DEC-2001 (first entry)
XX	
DE	Human taste receptor, hTIR3 cDNA coding sequence.

KV	Human; taste-cell-specific G protein-coupled receptor; hTIR3; drug;
KW	genetic modulation; pharmaceutical; taste sensation; food industry;
KM	chemosensory transduction; chromosome 1p36.2-1p36.33; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	CDS 1..2253
FT	/tag=
FT	/product= "Human taste receptor, hTIR3 protein"
PN	WO200166563-A2.
XX	
PD	13-SEP-2001.
XX	
PF	07-MAR-2001; 2001WO-US07265.
XX	
PR	07-MAR-2000; 2000US-0187546.
PR	07-APR-2000; 2000US-0195536.
PR	06-JUN-2000; 2000US-0209840.
PR	23-JUN-2000; 2000US-0214213.
PR	17-AUG-2000; 2000US-0226448.
PR	03-JAN-2001; 2001US-0255227.
XX	
PA	(SENO-) SENOMYX INC.
PI	Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L,
XX	WPI: 2001-582267/65.
DR	P-FSDB; AAIE10366.
XX	
PT	New mammalian taste-cell-specific G protein-coupled receptor
PT	polypeptides for identifying compounds that modulate taste signaling
XX	are useful in food, to modulate the sweet taste of foods or drugs -
PS	Claim 1; Page 72-73; 119pp; English.
XX	
CC	The invention relates to mammalian taste-cell-specific G protein-coupled
CC	receptors, TIR and their corresponding cDNA molecules. Taste receptors,
CC	TIR are useful for screening compounds which are used to activate or
CC	modulate chemosensory transduction, such as taste sensation. The
CC	identification and isolation of novel taste receptors and taste
CC	signalling molecules allow for new methods of chemical and genetic
CC	modulation of taste transduction pathways. The taste modulating
CC	compounds are useful in pharmaceuticals and food industries to improve
CC	the taste of a variety of consumer products, or to block undesirable
CC	tastes, e.g., in certain pharmaceuticals. TIR's are also useful in
CC	biochemical assay for identifying tastant (TIR) ligands having binding
CC	specificity for TIR involved in taste signalling. The present cDNA
CC	sequence is human taste-cell-specific G protein-coupled receptor, hTIR3
CC	coding sequence. Human TIR3 gene is localised on chromosome
CC	1p36.2-1p36.33.
XX	
SQ	Sequence 2553 BP; 379 A; 861 C; 838 G; 475 T; 0 other;
XX	
Query Match	99.1%; Score 2537; DB 22; Length 2553;
Best Local Similarity	99.8%; Pred. No. 0;
Matches 2553; Conservative	0; Mismatches 0; Indels 6; Gaps 1
OY	1 ATGCTGGGCCCTGCTGTCTCCTTGAGCCTCAACCTCTGTGGCTCTCTGCACCCCTGGAGCGGG 60
Dd	1 ATGCTGGGCCCTGCTGTCTCCTTGAGCCTCAACCTCTGTGGCTCTCTGCACCCCTGGAGCGGG 60
OY	61 GCCCATGTCGCTGCTGCACAGCAACTTGAAGAAGAGGGGAGTACTGCTGGGGGGCTG 120
Dd	61 GCCCATGTCGCTGCTGCACAGCAACTTGAAGAAGAGGGGAGTACTGCTGGGGGGCTG 120
OY	121 TTCCCCCTGGGGGAGCGCGAGGAGGCTGTCCGACGCGGACAGCGGCCAGACGCTT 180
Dd	121 TTCCCCCTGGGGGAGCGCGAGGAGGCTGTCCGACGCGGACAGCGGCCAGACGCTT 180
OY	181 GTGTGACACAGGTTTCTCTCAAAAGGCTGCTGTGGGACATGCGCATGAAAATGGCGCTG 240
Dd	181 GTGTGACACAGGTTTCTCTCAAAAGGCTGCTGTGGGACATGCGCATGAAAATGGCGCTG 240

Db 2461 CACCTGCCAGGTCTACCTGCTCATGCGGACGCCAGGGCTCAACACCCCGAGTTCTTC 2520
 QY 2488 CTGGGAGGGGGCCCTGGGGAGATGCCAAGGCAGATGACGGGAACACAGAAATCAGGGG 2547
 Db 2521 CTGGGAGGGGGCCCTGGGGAGATGCCAAGGCAGATGACGGGAACACAGAAATCAGGGG 2580
 QY 2548 AAACATGAGTGA 2559
 Db 2581 AAACATGAGTGA 2592

RESULT 7
 ABK16615
 ID ABK16615 standard; cDNA; 2739 BP.
 XX ABK16615;
 AC
 XX 14-MAR-2002 (first entry)
 DT
 XX Human G-coupled receptor (GCREC) cDNA, Seq ID No 24.
 DE
 XX Human; cytosolic; neuroprotective; immunosuppressant; nocrotropic;
 KW anti-inflammatory; anti-viral; gastrointestinal; cardiovascular;
 KW cerebroprotective; G-coupled receptor; cell proliferative disease;
 KW lymphoma; leukaemia; breast cancer; cirrhosis; neurological disorder;
 KW stroke; Alzheimer's disease; multiple sclerosis; mental retardation;
 KW cardiovascular disease; atherosclerosis; angina pectoris; indigestion;
 KW congestive heart failure; gastrointestinal disorder; dysphagia; AIDS;
 KW gastritis; autoimmune disorder; inflammatory disorder; Crohn's disease;
 KW systemic lupus erythematosus; metabolic disorder; diabetes; obesity;
 KW viral infection; herpesvirus; parvovirus;
 KW acquired immune deficiency syndrome; ss.
 KW
 XX Homo sapiens.
 OS
 XX NC0200190359-A2.
 PN
 XX 29-NOV-2001.
 PD
 XX
 XX 22-MAY-2001; 2001MO-US16833.
 PF
 XX 22-MAY-2000; 2000US-206222P.
 PR 25-MAY-2000; 2000US-207476P.
 PR 02-JUN-2000; 2000US-208834P.
 PR 02-JUN-2000; 2000US-208861P.
 PR 07-JUN-2000; 2000US-209868P.
 XX
 XX (INCYTE) INCYTE GENOMICS INC.
 PA
 XX Patterson C, Tribouley CM, Yao MG, Griffin JA, Thornton M, Lu Y;
 PI Kallik DA, Gandhi AR, Au-Yang J;
 XX
 XX WPI; 2002-106199/14.
 DR P-PSDB; AAU80493.
 XX
 PT New G-protein coupled receptors useful for treating or preventing cell
 PT proliferative (e.g. leukemia), neurological (e.g. stroke),
 PT cardiovascular or autoimmune/inflammatory disorders -
 PT
 XX
 PS Claim 5; Page 137-138; 148pp; English.
 XX
 CC The invention relates to a novel human G-coupled receptor (I). (I) and
 CC its corresponding polynucleotides are useful for diagnosing, treating or
 CC preventing cell proliferative diseases (e.g. lymphoma, leukaemia, breast
 CC cancer or cirrhosis), neurological disorders (e.g. stroke, Alzheimer's
 CC disease, multiple sclerosis or mental retardation), cardiovascular
 CC diseases (e.g. atherosclerosis, angina pectoris or congestive heart
 CC failure), gastrointestinal disorders (e.g. dysphagia, indigestion or
 CC gastritis), autoimmune/inflammatory disorders (e.g. AIDS, Crohn's disease
 CC or systemic lupus erythematosus) or metabolic disorders (e.g. diabetes
 CC or obesity), or viral infections (e.g. infection by herpesvirus or
 CC parvovirus). ABK16615-ABK16637 represent novel human G-coupled
 CC receptor coding sequences of the invention.

XX Sequence 2739 BP; 415 A; 951 C; 872 G; 501 T; 0 other;
 SQ
 Query Match 81.7%; Score 2090.6; DB 24; Length 2739;
 Best Local Similarity 91.4%; Pred. No. 0;
 Matches 2455; Conservative 0; Mismatches 19; Indels 213; Gaps 14;

QY 1 ATGCTGGGCCCCCTGCTGCTCTGCTGGGCTCAGCCCTCGGGCTCTCTGCAACCTCGGACGGG 60
 Db 1 ATGCTGGGCCCCCTGCTGCTCTGCTGGGCTCAGCCCTCGGGCTCTCTGCAACCTCGGACGGG 60
 QY 61 GCCCATTTGTGCTCTGTCACAGCAACTTAGATGTAAGGGGACTACGTCTGGGGGCTG 120
 Db 61 GCCCATTTGTGCTCTGTCACAGCAACTTAGATGTAAGGGGACTACGTCTGGGGGCTG 120
 QY 121 TTCCCTGGGGGAGGCGGAGAGAGGCTGGCTCCGACGGCGAGACGGCCAGAGGCT 180
 Db 121 TTCCCTGGGGGAGGCGGAGAGAGGCTGGCTCCGACGGCGAGACGGCCAGAGGCT 180
 QY 181 GTGTGCACAGGTTCTCCTCAAAAGGCTGCTGTGGGCACCTGAGCCATGAATGGCGTG 240
 Db 181 GTGTGCACAGGTTCTCCTCAAAAGGCTGCTGTGGGCACCTGAGCCATGAATGGCGTG 240
 QY 241 GAGGAGATCAACAACAAGTGGATCTGCTGGCGGCTGGCGCTGAGCTTCTT 300
 Db 241 GAGGAGATCAACAACAAGTGGATCTGCTGGCGGCTGGCGCTGAGCTTCTT 300
 QY 301 GATACGTGCTGCGAGCCCTGTTGGTGGCCATGAAAGCCCATGTTCTTGGCCAAAGCA 360
 Db 301 GATACGTGCTGCGAGCCCTGTTGGTGGCCATGAAAGCCCATGTTCTTGGCCAAAGCA 360
 QY 361 GGCAGCCGCGACATCGCCGCTACTGCAACTACAGCAAGTACAGCCCGTGTGCTGCT 420
 Db 361 GGCAGCCGCGACATCGCCGCTACTGCAACTACAGCAAGTACAGCCCGTGTGCTGCT 420
 QY 421 GTCATGGGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 Db 421 GTCATGGGGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 481 CTCATGCC-----CAGTCACTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 489
 Db 481 CTCATGCCCGGAGAGACCTTCCCTCTTCCGACCGGCGGAGAGAGAGAGAGAGAGAGAG 540
 QY 490 -----CAGTCACTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 516
 Db 541 GCCCTGTGTCAGAGAGATGCTCTTGGCCCTTGGAGGTGACGCTACGCTGCTGCTGCTGCT 600
 QY 517 CTGCTGAGCGCGGAGAGACCTTCCCTCTTCCGACCGGCGGAGAGAGAGAGAGAGAGAGAG 576
 Db 601 CTGCTGAGCGCGGAGAGACCTTCCCTCTTCCGACCGGCGGAGAGAGAGAGAGAGAGAGAG 660
 QY 577 CAGCTGACGGCGCGGAGAGCTCTGCAAGAGTTGGCTGGAATGGGTGGCCGCTG 636
 Db 661 CAGCTGACGGCGCGGAGAGCTCTGCAAGAGTTGGCTGGAATGGGTGGCCGCTG 720
 QY 637 GGCAGCAGCAGCAGTACGGCGGCGGAGAGCTTGTGAGCATTTTGTGGCCCTGCGCGCA 696
 Db 721 GGCAGCAGCAGCAGTACGGCGGCGGAGAGCTTGTGAGCATTTTGTGGCCCTGCGCGCA 779
 QY 697 CGGCGCATCTGCAATCGGCGAGAGAGGCTGGTGGCGGCTGGCGGATGATCGCGG 756
 Db 780 CGGCGCATCTGCAATCGGCGAGAGAGGCTGGTGGCGGCTGGCGGATGATCGCGG 822
 QY 757 CTGGGAGAGGTGACAGAGTCTCTGACAGGTGAACAGAGACAGTGTGCTGCTG 816
 Db 823 -----GTGACAGAGTCTCTGACAGGTGAACAGAGACAGTGTGCTGCTG 873
 QY 817 CTGTGCTCTCTGTCACGCGCGCCACGCTTCTTCACTACATACAGATACAGAGCTC 876
 Db 874 CTGTGCTCTCTGTCACGCGCGCCACGCTTCTTCACTACATACAGATACAGAGCTC 933
 QY 877 TCGCCCAAGTGTGTGGTGGCAGGAGAGCTGGTGAACCTTGAACCTGTGATGAGGCTG 936
 Db 877 TCGCCCAAGTGTGTGGTGGCAGGAGAGCTGGTGAACCTTGAACCTGTGATGAGGCTG 936

D	b	934	TCGCCCAAGGTTGGGTGGCCACGAGAGGCTGGCTACCTCTTGACCTGTGCTATGGGGCTG	993
O	y	937	CCCGCATGGCCCCAGATGGGCACGGTCGTTGGCTTCTCAGAGGGAGTGGCCACGGCTGCAC	996
D	b	994	CCCGGCATGGGCCAGATGGGGCACGGTGTGGCTTCTCTCAAGAGGGAGTCCACGGCTGCAC	1053
O	y	997	GAGTTCCTCCCGATGTCGTGAAGAGCGCACCTGGCCCTGGCCACCAGACCCGGCTTCTGTCTCT	1056
D	b	1054	GAGTTCCTCCCGATGTCGTGAAGAGCGCACCTGGCCCTGGCCACCAGACCCGGCTTCTGTCTCT	1113
O	y	1057	GCCCTGGGGCGAGAGGAGACAGGGGTCGTGAAGAGGAGCATGTGGTGGGGCAGAGCGTGGCCGAG	1116
D	b	1114	GCCCTGGGGCGAGAGGAGACAGGGGTCGTGAAGAGGAGCATGTGGTGGGGCAGAGCGTGGCCGAG	1173
O	y	1117	TGTGATCTGCATCACTGTGCAGAAACGTGTAGCGCAGGGGCTAAATCACACACAGACGTTCTCT	1176
D	b	1174	TGTGATCTGCATCACTGTGCAGAAACGTGTAGCGCAGGGGCTAAATCACACACAGACGTTCTCT	1196
O	y	1177	GTTTACGACGCTGTGTATAGCGTGGCCACAGGCCCTGCACACACTTCTATGTGCACGGCC	1236
D	b	1197	-----CCGTGGCCAGGGCCCTGTGCACACACTTCTATGTGCACGGCC	1236
O	y	1237	TCAGAGCTGCCCCGGCGAGAGACCCCGTGAAGACCCCTGGCACCTCTCTGGAAGAACTGTACAC	1296
D	b	1237	TCAGAGCTGCCCCGGCGAGAGACCCCGTGAAGACCCCTGGCACCTCTCTGGAAGAACTGTACAC	1296
O	y	1297	CTGACCTTTCACAGTGGGGCGGGCTGGCCGTTTGCAGCAGCAGCGGAAACGTGTGACATG	1356
D	b	1297	CTGACCTTTCACAGTGGGGCGGGCTGGCCGTTTGCAGCAGCAGCGGAAACGTGTGACATG	1356
O	y	1357	GAGTACACACTGAAAGCTGTGGGTGTGGCAGAGGCTCAGTGCCTCAGAGCTCCAGAGCTGGGC	1416
D	b	1357	GAGTACACACTGAAAGCTGTGGGTGTGGCAGAGGCTCAGTGCCTCAGAGCTCCAGAGCTGGGC	1416
O	y	1417	AGGTTCAACGGCAGACCTCAGAGACAGAGCGCTGAATATCGCTGGACACAGCTGTACACAC	1476
D	b	1417	AGGTTCAACGGCAGACCTCAGAGACAGAGCGCTGAATATCGCTGGACACAGCTGTACACAC	1476
O	y	1477	-----CAGAACCCGTTGTCCCGGTGTCCGCGG	1503
D	b	1477	CAGCCGAGACAGAGCCACAGACCCCAAGGGCTGTGGCGACAGAACCCGTTGTCCGCGGTGCGGG	1536
O	y	1504	CAGTGCAGAGAGGGCCAGAGTGGCCGCGGGTTCAMAGGGTTCACACTCTGCTGTACGACTGT	1553
D	b	1537	CAGTGCAGAGAGGGCCAGAGTGGCCGCGGGTTCAMAGGGTTCACACTCTGCTGTACGACTGT	1596
O	y	1564	GTTGGATCGCGAGGGGGGCGACCTACACGGGCAAAACCCAGAGAGATGGCTCGACACTTTTGT	1623
D	b	1537	GTTGGATCGCGAGGGGGGCGACCTACACGGGCAAAACCCAGAGAGATGGCTCGACACTTTTGT	1656
O	y	1624	GCGCAGAGATGAGTGTGTCCCGGAGACAGACACGCTGTTCGCGCGCAGAGTCTGGTTC	1683
D	b	1657	GCGCAGAGATGAGTGTGTCCCGGAGACAGACACGCTGTTCGCGCGCAGAGTCTGGTTC	1716
O	y	1684	CTGGCATGGGGCGAGCCGCGCTGTGCTGTGCTGTCTCTGTCTGTAGCCCTGGCGCTGGCC	1743
D	b	1717	CTGGCATGGGGCGAGCCGCGCTGTGCTGTGCTGTCTCTGTCTGTAGCCCTGGCGCTGGCC	1776
O	y	1744	CTTGGCGAGGCTGTGTTGGGGGCGTTTCGTTTCACACATCGGGGAAGCGCCACTGGTTTACGGCC	1803
D	b	1777	CTTGGCGAGGCTGTGTTGGGGGCGTTTCGTTTCACACATCGGGGAAGCGCCACTGGTTTACGGCC	1836
O	y	1804	TCGGGGGGGGCCCTGGGCTCTTGTGGCTGTGGTGTGCTGTGGGCTGGTCTGAGCTTCAGCGTC	1863
D	b	1837	TCGGGGGGGGCCCTGGGCTCTTGTGGCTGTGGTGTGCTGTGGGCTGGTCTGAGCTTCAGCGTC	1896
O	y	1864	CTCTGTTCCCTGTGGCCAGCCCAAGCCCTTGCCTGCAATGCTGGCCAGACAGCCCTTGTCCAC	1923
D	b	1897	CTCTGTTCCCTGTGGCCAGCCCAAGCCCTTGCCTGCAATGCTGGCCAGACAGCCCTTGTCCAC	1956
O	y	1924	CTCCCCGCTCAAGGGGTCTCTGTAGACACAATCTTCTGTCAAGGGCGCGAGATCTTGTGTGAG	1983
D	b	1957	CTCCCCGCTCAAGGGGTCTCTGTAGACACAATCTTCTGTGTAGGGCGCGAGATCTTGTGTGAG	2016

QY	1984	TCGAACGCGCTCTGAGACTGTGGGAGACGCCGCTGAGTGGCTGGCTGGGGGCGCCGAGGCC	204.3
Db	2017	TCAGAACTGCTCTGAGCTTGGGACAGACGGCTGAGTGGCTGGCTGGGGGCGCCGAGGCC	207.6
QY	2044	TGGCTGGTGGTGGCTGGCTGGGCCATGCTGGTGGAGAGTGGCATGTGGACCTGGACTGGTGTG	210.03
Db	2077	TGGCTGGTGGTGGCTGGCTGGGCCATGCTGGTGGAGAGTGGCATGTGGACCTGGACTGGTGTG	213.6
QY	2104	GCTTTCGCCGCGAGGTGGTGAC - GGACTGGACATGCTGCCACGGAGGCGCTGGTGCA	216.2
Db	2137	GCTTTCGCCGCGAGGTGGTGAGTGGAGTGGACTGGACCTGGACATGCTGCCACGGAGGCGCTGGTGCA	219.6
QY	2153	CTGGCCGACACAGCTCTCTGGGCTCAGCTTCGGGCTCAGGCGACGCCACCAATGCAAGCTGGC	222.22
Db	2197	CTGGCCGACACAGCTCTCTGGGCTCAGCTTCGGGCTCAGGCGACGCCACCAATGCAAGCTGGC	225.55
QY	2223	CTTTCCTGAGCTTC - CTGGAGCATTTTCTGGTGGAGGACCGCGGGCTGCTCAACCGTG	228.1
Db	2256	CTTTCCTGAGCTTCTGACTGTGGGCACTTCTGGTGGGAGCCAGCGCGGCTGCTCAACCGTG	231.5
QY	2282	CC - - CGTGGCTTACCTTTGGCCATGCTGGCTGGCTACTTTCATCAC - CTGGGCTTC - CTTTGT	233.6
Db	2316	CCAGGTGGCCCTCACACATTTGGCATGTGGCTGGCTACTTTCATCACCGGGCTTCACTTTGT	237.5
QY	2337	GCCCTCTCTGGC - CAATGGTCAAGTGGT - - CTCACAGGCCCGCGTGCAGATGGGGGCGCT	239.93
Db	2376	GCCCTCTCTGGCAGTAATGTGAGGTGGTCACTCAGGCCCCAGCGCTGCAGATGGGGGCGCT	243.55
QY	2394	CTGTCTGTGTGTCCTGGGCAATCTGTGGCTTCCACCTGCGCAGGTGTTACTGTCTCAT	245.3
Db	2436	CTGTCTGTGTGTCCTGGGCAATCTGTGGCTTCCACCTGCGCAGGTGTTACTGTCTCAT	249.93
QY	2454	GCGGACGCAAGGCTCAACACCCCCGAGTTCTTCTGGGAGGGGGCCCTGGGGATGCC - C	251.2
Db	2496	GCGGACGCAAGGCTCAACACCCCCGAGTTCTTCTGGGAGGGGGCCCTGGGGATGCCAC	255.55
QY	2513	AAGGCGAAGATGACGGGAAACAGAGGAATCAGGGGAAACATGAGTGA	255.9
Db	2556	AAGGCGAAGATGACGGGAAACAGAGGAATCAGGGGAAACATGAGTGA	260.2
RESULT 8			
AL38463			
ID	AL38463	standard; DNA; 8001 bp.	
XX	XX	AL38463;	
AC	XX		
XX	DT	15-AUG-2002 (first entry)	
XX	DE	Genomic DNA encoding a human G-protein coupled receptor (GPCR).	
XX	XX	G-protein coupled receptor; GPCR; gene chip; human; immune response;	
KM	KW	chromosomal position; transgenic animal; gene therapy; gene; ds..	
XX	OS	Homo sapiens.	
XX	XX		
PH	Key	Location/Qualifiers	
FT	CDS	2118..5236	
FT	FT	/*tag= a	
FT	FT	/product= "Human G-protein coupled receptor"	
FT	FT	/note= "This coding sequence contains 5 introns"	
FT	FT	2118..2308	
FT	FT	/*tag= b	
FT	FT	/number= 1	
FT	FT	2309..2409	
FT	FT	/*tag= c	
FT	FT	/number= 1	
FT	FT	2410..2710	
FT	FT	/*tag= d	
FT	FT	/number= 2	
FT	FT	2711..2795	
FT	FT	/*tag= e	
FT	FT	Intron	

FT	/number- 2
FT	2796..3578
FT	/*tag= f
FT	/number- 3
FT	3579..3692
FT	/*tag= g
FT	/number- 3
FT	3693..3896
FT	/*tag= h
FT	/number- 4
FT	3897..4030
FT	/*tag= i
FT	/number- 4
FT	4031..4151
FT	/*tag= j
FT	/number- 5
FT	4152..4427
FT	/*tag= k
FT	/number- 5
FT	4278..5236
FT	/*tag= l
FT	/number- 6
FT	replaced (406,C)
FT	/*tag= m
FT	replaced (852,C)
FT	/*tag= n
FT	replaced (897,A)
FT	/*tag= o
FT	replaced (1433,T)
FT	/*tag= p
FT	replaced (5845,C)
FT	/*tag= q
FT	replaced (7028,A)
FT	/*tag= r
XX	
PN	WO200230981-AI.
PD	18-APR-2002.
XX	
PF	13-MAR-2001; 2001WO-US07832.
XX	
PR	10-OCT-2000; 2000US-0684393.
XX	
PA	(PEKE) PE CORP NT.
XX	
PI	WeI M, Zhong W, Ketchum KA, Difrancesco V, Beasley EM;
DR	WPI; 2002-444173/47.
XX	P-PSDB; AAO21501.
PT	Novel G protein coupled receptor, useful for raising antibodies, to
FT	elicit immune response, and as a reagent in assays designed to
XX	quantitatively determine levels of protein in biological samples -
SS	Disclosure; Fig 3; 82pp; English.
CC	
XX	
CC	The invention relates to an isolated G-protein coupled receptor (GPCR)
CC	polypeptide, comprising an 852 residue amino acid sequence, given in the
CC	specification, an allelic variant or ortholog of the protein, or a
CC	fragment comprising at least 10 contiguous amino acids of the protein.
CC	GPCR is useful for identifying a modulator of GPCR and an agent that
CC	binds to GPCR. GPCR and a gene chip comprising GPCR are useful as models
CC	for the development of human therapeutic agents. GPCR is useful for
CC	raising antibodies, to elicit immune response, as a reagent in assays
CC	designed to quantitatively determine levels of protein in biological
CC	samples, and as markers for tissues in which the corresponding protein
CC	is preferentially expressed. A gene chip containing GPCR is also useful
CC	as a probe for determining the chromosomal positions of nucleic acid
CC	molecules by means of in situ hybridisation. In making vectors containing
CC	the gene regulatory regions of a gene chip containing GPCR, for designing
CC	ribozymes, in making vectors that express GPCR, and for constructing host
CC	cells and transgenic animals expressing nucleic acid molecules and
CC	peptides. A host cell containing GPCR is useful for conducting cell-based

CC	assays involving GPCR protein or its fragments, and for identifying
CC	protein mutants in which these function is affected. The polynucleotide
CC	encoding PCR can be used to treat disorders by gene therapy. This
CC	polynucleotide sequence represents The genomic DNA encoding the human G-
CC	protein coupled receptor of the invention.
xx	
xx	
Sequence 8001 BP; 1291 A; 2630 C; 2623 G; 1457 T; 0 other;	
Query Match	76.1%; Score 1947.4; DB 24; Length 8001;
Best Local Similarity	82.0%; Pred. No. 0;
Matches 2558; Conservative	0; Mismatches 1; Indels 560; Gaps
1	AGCTGAGGCGCTGCTGCTCTGAGCCCTGAGCCCTCTGAGGCTCTCTGACACCTGGAGCGG 60
2118	ATGCTGAGGCGCTGCTGCTCTGAGCCCTGAGCCCTCTGAGGCTCTCTGACACCTGGAGCGG 2177
61	GCCCATTTGCTGTACACGCACTTAGATGAAGGGGACTACTGCTGGGGGGCTG 120
2178	GCCCATTTGCTGTACACGCACTTAGATGAAGGGGACTACTGCTGGGGGGCTG 2237
121	TTCCCCCTGGGGGAGGCGGAGGCGCTGGCCCTCCGACGCGGACACGGGCCAGAGCCCT 180
2238	TTCCCCCTGGGGGAGGCGGAGGCGCTGGCCCTCCGACGCGGACACGGGCCAGAGCCCT 2297
181	GTGTGCA----- 187
2298	GTGTGACACAGGTACAGAGGTGGAGCGGCGTGGTCGGGGTCAAGGTGACAGGTCGTGGG 2357
188	-----CAGAGTTCCTCT 199
2358	GTGCTCTGAGCTGGGGCGGAGGTGGCCATCGGGTTCTGTGTGGCCCCCAGAGTTCCTCT 2417
200	CAAGGCGCTGCTGTGGGCACTGGCCATGAATAATGGCGTGGAGAGATCAACACAAGT 259
2418	CAAGGCGCTGCTGTGGGCACTGGCCATGAATAATGGCGTGGAGAGATCAACACAAGT 2477
260	CGAGTCTGCTGCCGGGCTGCGCTGGGCTACGACCTCTTTGATACGTGCTCGAGCCTG 319
2478	CGAGTCTGCTGCCGGGCTGCGCTGGGCTACGACCTCTTTGATACGTGCTCGAGCCTG 2537
320	TGGTGCCATGAAGGCCAGGCTCATGTTCTGTCGGCAAGGAGGCGAGCGGCACATCGGCG 379
2538	TGGTGCCATGAAGGCCAGGCTCATGTTCTGTCGGCAAGGAGGCGAGCGGCACATCGGCG 2597
380	CCTACTGCAATACAGCAGATACAGCCCGGTGTCTGGCTGTCTCATCGGGCCCCACTCGT 439
2598	CCTACTGCAATACAGCAGATACAGCCCGGTGTCTGGCTGTCTCATCGGGCCCCACTCGT 2657
440	CAGACTCGCCATGGTCAACGGCAGTCTTAGTTCCTCTCATGGCC----- 489
2658	CAGACTCGCCATGGTCAACGGCAGTCTTAGTTCCTCTCATGGCC----- 2717
490	----- 489
2718	CCCCCACCACATACCACCCACCACGCCCTGCGCGGTGGAGGCCCTGTGCAGAGAT 2777
490	-----CAGTCACTACGCTGTACATGACATGACCTGACGCGCCCGGAG 534
2778	GCTCTTGGCCCTTACAGGTCACTACGCTGTGTACATGACATGACCTGACGCGCCCGGAG 2837
535	ACCTTCCCTCTTTCGCGACCGTGGCCAGCGACCGTGGACCTGACGAGCGCGCGCG 594
2838	ACCTTCCCTCTTTCGCGACCGTGGCCAGCGACCGTGGACCTGACGAGCGCGCGCGCG 2897
595	GAGTCTGTCAGAGATTGGCTGGAATGGTGGCGGCCCTGGGCGAGCGAGCAGAGTAC 654
2898	GAGTCTGTCAGAGATTGGCTGGAATGGTGGCGGCCCTGGGCGAGCGAGCAGAGTAC 2957
655	GCCCGGAGGCGCTTACGATCTTCTCGGCCCTCTGGCGCGGACAGGGGATCTGCATCGCG 714
2958	GCCCGGAGGCGCTTACGATCTTCTCGGCCCTCTGGCGCGGACAGGGGATCTGCATCGCG 3017
715	CACGAGGCGCTGGTCCCTGCGCGTCCGATGACTCGCGGCTGGGGGAAGGTGAGAGAC 774

RESULT 9
 AAD17517 standard; DNA: 3563 BP.
 ID AAD17517;
 AC AAD17517;
 XX
 XX
 DT 10-DEC-2001 (first entry)
 XX
 XX Human taste receptor, hTIR3 full-length genomic DNA.
 DE
 KW Human: taste-cell-specific G protein-coupled receptor; TIR3; drug;
 KW genetic modulation; pharmaceutical; taste sensation; food industry;
 KW chemosensory transduction; chromosome 1p36.2-1p36.33; ds.
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3381
 FT /tag- a
 FT /product- "Human taste receptor, hTIR3 protein"
 FT 1..210
 FT /tag- b
 FT /number- 1
 FT /cons_splice- (5'site:NO, 3'site:NO)
 FT 211..401
 FT /tag- c
 FT /number- 1
 FT 402..502
 FT /tag- d
 FT /number- 2
 FT 503..802
 FT /tag- e
 FT /number- 2
 FT 803..945
 FT /tag- f
 FT /number- 3
 FT 946..1723
 FT /tag- g
 FT /number- 3
 FT 1724..1837
 FT /tag- h
 FT /number- 4
 FT 1838..2041
 FT /tag- i
 FT /number- 4
 FT 2042..2175
 FT /tag- j
 FT /number- 5
 FT 2176..2296
 FT /tag- k
 FT /number- 5
 FT 2297..2422
 FT /tag- l
 FT /number- 6
 FT 2423..3381
 FT /tag- m
 FT /number- 6
 FT 3382..3563
 FT /tag- n
 FT /number- 7
 FT /cons_splice- (5'site:NO, 3'site:NO)
 FT
 XX
 XX
 PN MO200166563-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 07-MAR-2001; 2001MO-US07265.
 XX
 PR 07-MAR-2000; 2000US-0187546.
 PR 07-APR-2000; 2000US-0195536.
 PR 06-JUN-2000; 2000US-0209840.

PR 23-JUN-2000; 2000US-0214213.
 PR 17-AUG-2000; 2000US-0226448.
 PR 03-JAN-2001; 2001US-0259227.
 XX
 XX
 PA (SENO-) SENOMYX INC.
 PI Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
 PI WPI; 2001-582267/65.
 DR P-PSDB; AAE10366.
 XX
 XX
 PT New mammalian taste-cell-specific G protein-coupled receptor
 PT polypeptides for identifying compounds that modulate taste signaling
 PT are useful in food, to modulate the sweet taste of foods or drugs
 PS Claim 1; Page 71-72; 119pp; English.
 XX
 CC The invention relates to mammalian taste-cell-specific G protein-coupled
 CC receptors, TIR and their corresponding cDNA molecules. Taste receptors,
 CC TIR are useful for screening compounds which are used to activate or
 CC modulate chemosensory transduction, such as taste sensation. The
 CC identification and isolation of novel taste receptors and taste
 CC signalling molecules allow for new methods of chemical and genetic
 CC modulation of taste transduction pathways. The taste modulating
 CC compounds are useful in pharmaceuticals and food industries to improve
 CC the taste of a variety of consumer products, or to block undesirable
 CC tastes, e.g., in certain pharmaceuticals. TIRs are also useful in
 CC biochemical assay for identifying tastant (TIR) ligands having binding
 CC specificity for TIR involved in taste signalling. The present sequence is
 CC human taste-cell-specific G protein-coupled receptor, hTIR3 full-length
 CC genomic DNA. Human TIR3 gene is localised on chromosome 1p36.2-1p36.33.
 XX
 XX
 SQ Sequence 3563 BP; 545 A; 1214 C; 1166 G; 638 T; 0 other;
 Query Match 74.1%; Score 1897; DB 22; Length 3563;
 Best Local Similarity 80.7%; Pred. No. 5, 3e-313;
 Matches 2559; Conservative 0; Mismatches 0; Indels 612; Gaps 5;
 QY 1 ATGCTGGGCCCTGCTGCTGCTGCGGCTCAGCCTCTCTGTCACACCTGGAGCGGG 60
 DB 211 ATGCTGGGCCCTGCTGCTGCTGCGGCTCAGCCTCTCTGTCACACCTGGAGCGGG 270
 QY 61 GCCCATTTGCTGCTGTACAGCACTTAGATGAAGGGGACATGCTGCGGGGGCTG 120
 DB 271 GCCCATTTGCTGCTGTACAGCACTTAGATGAAGGGGACATGCTGCGGGGGCTG 330
 QY 121 TTCCTGCGGCGAGGCGGAGAGCTGCTCGGACGCGGACAGCGCCAGACGCT 180
 DB 331 TTCCTGCGGCGAGGCGGAGAGCTGCTCGGACGCGGACAGCGCCAGACGCT 390
 QY 181 GTGTGCA----- 187
 DB 391 GTGTGCAAGTAGAGAGGTGGAGCGGCTGGGTGCGAGGTGACAGAGTGTGG 450
 QY 188 -----CCAGGTCTCTCT 199
 DB 451 GTGCTCTGAGCTGGGGCGAGGTGGCCATCTGCGGTTCTGTGTGGCCCGAGGTCTCT 510
 QY 200 CAAACGGCTGCTCTGGGCACTGGCCATGAATAATGGCGGTGAGAGATCAACAACAGT 259
 DB 511 CAAACGGCTGCTCTGGGCACTGGCCATGAATAATGGCGGTGAGAGATCAACAACAGT 570
 QY 260 CGGATCTGTGCGCGGGCTCGGCTGGGCTACGACCTCTTTGATACGTGCTCGGAGCTG 319
 DB 571 CGGATCTGTGCGCGGGCTCGGCTGGGCTACGACCTCTTTGATACGTGCTCGGAGCTG 630
 QY 320 TGGTGGCCATGAAGCCAGCCTCATGTTCTGGCCAGAGGAGCGCGGACATGCGCG 379
 DB 631 TGGTGGCCATGAAGCCAGCCTCATGTTCTGGCCAGAGGAGCGCGGACATGCGCG 630
 QY 380 CTTACTGCACTACAGCACTACAGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 439
 DB 691 CTTACTGCACTACAGCACTACAGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 750


```

Db 2911 ACCGGAACCTGGGCTTCCCGGAGGTGTGACGGATGGCATCTGCTGCCAGC 2970
QY 2149 GAGGCGCTGGTGCACCTGGCCACAGCGCTCTGGGTCAGCTTGGGCTTGGCCAGCCACC 2208
Db 2971 GAGGCGCTGGTGCACCTGGCCACAGCGCTCTGGGTCAGCTTGGGCTTGGCCAGCCACC 3030
QY 2209 AATGCCAAGCTGGGCTTCTGCTTCTGCTTCTGAGGCACTTCTGAGGAGCCAGCCGGGC 2268
Db 3031 AATGCCAAGCTGGGCTTCTGCTTCTGCTTCTGAGGCACTTCTGAGGAGCCAGCCGGGC 3090
QY 2269 TGTACACACCTGGGCTTCTGCTTCTGCTTCTGAGGCACTTCTGAGGAGCCAGCCGGGC 2328
Db 3091 TGTACACACCTGGGCTTCTGCTTCTGCTTCTGAGGCACTTCTGAGGAGCCAGCCGGGC 3150
QY 2329 TCTTTTGGCCCTCTCTGGCCAAATGTGAGTGTCTCTCAGAGCCGCGCTGAGATGGGC 2388
Db 3151 TCTTTTGGCCCTCTCTGGCCAAATGTGAGTGTCTCTCAGAGCCGCGCTGAGATGGGC 3210
QY 2389 GCGCTCTGCTCTGTCTGTCTGGGCACTCTGCTCTCACTGCCAGGTGTACCTG 2448
Db 3211 GCGCTCTGCTCTGTCTGTCTGGGCACTCTGCTCTCACTGCCAGGTGTACCTG 3270
QY 2449 CTCATGGGCGACGAGGCTTAACACCCCGAGTTCTTCTGGGAGGGGCGCTGGGGAT 2508
Db 3271 CTCATGGGCGACGAGGCTTAACACCCCGAGTTCTTCTGGGAGGGGCGCTGGGGAT 3330
QY 2509 GCGCCAGGCGCAGAAATGACGAGGAGACAGAAATCAGGGAATCAGATGA 2559
Db 3331 GCGCCAGGCGCAGAAATGACGAGGAGACAGAAATCAGGGAATCAGATGA 3381

RESULT 10
AADI7508
ID AADI7508 standard; DNA: 2687 BP.
XX
AC AADI7508;
XX
DT 10-DEC-2001 (first entry)
XX
DE Human taste receptor, hTIR3 genomic DNA - 3' contig.
XX
KM Human: taste-cell-specific G protein-coupled receptor; hTIR3; drug;
KM genetic modulation; pharmaceutical; taste sensation; food industry;
KM chemosensory transduction; chromosome 1p36.2-1p36.33; 68.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT 1..69
FT /tag= a
FT /number= 1
FT /cons_splice= (5'site:NO, 3'site:YES)
FT exon /partial
FT 70..847
FT /tag= b
FT /number= 1
FT /tag= c
FT 848..961
FT /tag= c
FT /number= 2
FT exon /tag= d
FT 962..1165
FT /tag= d
FT /number= 2
FT Intron /tag= e
FT 1166..1299
FT /tag= e
FT /number= 3
FT exon /tag= f
FT 1300..1420
FT /tag= f
FT /number= 3
FT Intron /tag= g
FT 1421..1546
FT /tag= g
FT /number= 4
FT exon /tag= h
FT 1547..2505

```

```

FT /tag= h
FT /number= 4
FT 2506..2687
FT Intron /tag= i
FT /number= 5
FT /cons_splice= (5'site:NO, 3'site:NO)
XX
PN W020016563-A2.
XX
PD 13-SEP-2001.
XX
PF 07-MAR-2001; 2001MO-US07265.
XX
PR 07-MAR-2000; 2000US-0187546.
PR 07-APR-2000; 2000US-0195536.
PR 06-JUN-2000; 2000US-0209840.
PR 23-JUN-2000; 2000US-0214213.
PR 17-AUG-2000; 2000US-0226448.
PR 03-JAN-2001; 2001US-0259227.
XX
PA (SENO-) SENOMYX INC.
XX
PI Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
XX WPI; 2001-562267/65.
XX
DR WPI; 2001-562267/65.
XX
PT New mammalian taste-cell-specific G protein-coupled receptor
PT polypeptides for identifying compounds that modulate taste signaling
PT are useful in food, to modulate the sweet taste of foods or drugs -
XX
PS Claim 133; Page 69-71; 119pp; English.
XX
XX
CC The invention relates to mammalian taste-cell-specific G protein-coupled
CC receptors, TIR and their corresponding cDNA molecules. Taste receptors,
CC TIR are useful for screening compounds which are used to activate or
CC modulate chemosensory transduction, such as taste sensation. The
CC identification and isolation of novel taste receptors and taste
CC signalling molecules allow for new methods of chemical and genetic
CC modulation of taste transduction pathways. The taste modulating
CC compounds are useful in pharmaceuticals and food industries to improve
CC the taste of a variety of consumer products, or to block undesirable
CC tastes, e.g., in certain pharmaceuticals. TIRs are also useful in
CC biochemical assay for identifying tastant (TIR) ligands having binding
CC specificity for TIR involved in taste signalling. The present sequence is
CC human taste-cell-specific G protein-coupled receptor, hTIR3 genomic DNA -
CC 3' contig. Human TIR3 gene is localised on chromosome 1p36.2-1p36.33.
XX
SQ Sequence 2687 BP; 403 A; 920 C; 885 G; 479 T; 0 other;
XX
Query Match 65.18; Score 1666; DB 22; Length 2687;
Best Local Similarity 84.7%; Pred. No. 7.9e-274;
Matches 2070; Conservative 0; Mismatches 0; Indels 374; Gaps 3;
QY 490 CAGGTGACCTAGGCTGTAGATGAGCTGTGAGCCCGGAGGACCTCCCTCTTC 549
Db 62 CAGGTGACCTAGGCTGTAGATGAGCTGTGAGCCCGGAGGACCTCCCTCTTC 121
QY 550 TTCCGCAACGTCGCCAGCAGACGCTGTGAGTGAACGCGCGCGGAGCTGTGACAGAG 609
Db 122 TTCCGCAACGTCGCCAGCAGACGCTGTGAGTGAACGCGCGCGGAGCTGTGACAGAG 181
QY 610 TTGCGCTGGAACCTGGTGGCGCCCTTGGGACAGCAGACAGTACGGCCGGCAGGGCCTG 669
Db 182 TTGCGCTGGAACCTGGTGGCGCCCTTGGGACAGCAGCAGAGTACGGCCGGCAGGGCCTG 241
QY 670 AGCATCTTCTCGGCGCTGGCGCGGAGCGGAGCTGTGATCGGCGAGAGGGCTGGTG 729
Db 242 AGCATCTTCTCGGCGCTGGCGCGGAGCGGAGCTGTGATCGGCGAGAGGGCTGGTG 301
QY 730 CCGCTGCGCGCTGGCGGAGTGCAGTGCAGTGGGAGGTGACAGAGCTGTGACAGAGTG 789
Db 302 CCGCTGCGCGCTGGCGGAGTGCAGTGCAGTGGGAGGTGACAGAGCTGTGACAGAGTG 361

```

QY	790	AACAGAGCAGCTGCAGAGTGGTGTCTGTCTGTCCCTCCGTGCACAGCCGCCACGCCCTC	849
Db	362	AACAGAGCAGCTGCAGAGTGGTGTCTGTCTGTCCCTCCGTGCACAGCCGCCACGCCCTC	421
QY	850	TTCAACTACAGCATACAGACAGAGCTCTGGGCCAAGGTGTGGGTGGCCAGAGAGGCTGG	909
Db	422	TTCAACTACAGCATACAGACAGAGCTCTGGGCCAAGGTGTGGGTGGCCAGAGAGGCTGG	481
QY	910	CTGACCTCTACCTGGTGCATGAGGGGTGCCCGGCATGAGCCAGATGGGCAGGCTTGGC	969
Db	482	CTGACCTCTACCTGGTGCATGAGGGGTGCCCGGCATGAGCCAGATGGGCAGGCTTGGC	541
QY	970	TTCTCTCAGAGGGGGTGGCCAGCTGCACAGATTCCCCAGATAGTAAAGCAGCAGCTGGCC	1029
Db	542	TTCTCTCAGAGGGGGTGGCCAGCTGCACAGATTCCCCAGATAGTAAAGCAGCAGCTGGCC	601
QY	1030	CTGGCCACCCAGCCCGGCTTCTCTCTGTGCCCTGGGCGAGAGGGAGCAGAGGTCTGGAGAG	1089
Db	602	CTGGCCACCCAGCCCGGCTTCTCTCTGTGCCCTGGGCGAGAGGGAGCAGAGGTCTGGAGAG	661
QY	1090	GACGTGGTGGGGCAGCGCTCCCGCAGTGTGACTGTGATACAGCTCAGAAAGTGAAGCGCA	1149
Db	662	GACGTGGTGGGGCAGCGCTCCCGCAGTGTGACTGTGATACAGCTCAGAAAGTGAAGCGCA	721
QY	1150	GGGCTAAATCACCACACAGAGAGTCTCTGTCTAGCGAGCTGTGTATAGGTGGCCAGAGCC	1209
Db	722	GGGCTAAATCACCACACAGAGAGTCTCTGTCTAGCGAGCTGTGTATAGGTGGCCAGAGCC	781
QY	1210	CTGACACACACTTTTACGTGACAAAGCCTTAAGCTGCCCGCAGAGACCCCGTGAAGCC	1269
Db	782	CTGACACACACTTTTACGTGACAAAGCCTTAAGCTGCCCGCAGAGACCCCGTGAAGCC	841
QY	1270	TG-----	1271
Db	842	TGGAGGTGAGCCCGGAGATGGGGGTGTGTCTCTCTGCATGTGCCAGCCACACAGG	901
QY	1272	-----GCAG	1275
Db	902	CACGAGCACCAGCCCTGAGAGTGGCTGGCGGAGCTCAGCCCGCTGCCCGGCCGAG	961
QY	1276	CTCTTGGAGAACTGTACAACTGCACCTTCCACAGTGGGGGCTCCGCTGGGTTGCAC	1335
Db	962	CTCTTGGAGAACTGTACAACTGCACCTTCCACAGTGGGGGCTCCGCTGGGTTGCAC	1021
QY	1336	AGCAGCGGAAGCGTGGATGAGTACGACCTGAAGCTGTGGGTGGCAGAGGGCTCAAGTG	1395
Db	1022	AGCAGCGGAAGCGTGGATGAGTACGACCTGAAGCTGTGGGTGGCAGAGGGCTCAAGTG	1081
QY	1396	CCGAGGCTCCAGCAGCGTGGGCTCAACGGCAGCTTCAGACACAGAGCGCTGTAAGATC	1455
Db	1082	CCGAGGCTCCAGCAGCGTGGGCTCAACGGCAGCTTCAGACACAGAGCGCTGTAAGATC	1141
QY	1456	CGCTGGCAGCAGTCTGACAAAC-----	1476
Db	1142	CGCTGGCAGCAGTCTGACAAACAGGTGAGGTGAGGTGGGTGGCCAGCGTGGCCCGTG	1201
QY	1477	-----	1476
Db	1202	TAGCCCCCGGCGAGGGCGCAGCTGGGGGTGGGGGCGCTTCACTCTCCGTGGGCAATG	1261
QY	1477	-----CAGAGCCCGCTGTCCGAGTGGCTGC	1501
Db	1262	CCGAGCGGAGCAGAGCGCAGACCCCAAGGCTGTGGCGAGAGAGCGGCTGTCCGGGTGGC	1321
QY	1502	GGCAGTCCCAAGAGGGCCAGGTGCGCCGGGTCAAGGGGTTCACATCTGTGTACGACT	1561
Db	1322	GGCAGTCCCAAGAGGGCCAGGTGCGCCGGGTCAAGGGGTTCACATCTGTGTACGACT	1381
QY	1562	GTTGGAGCTCGAGGGGGGAGGTATACCGGCAAAACC-----	1597
Db	1382	GTTGGAGCTCGAGGGGGGAGGTATACCGGCAAAACCAGGTAGCCCGCTTCCGGGCAAG	1441
QY	1598	-----	1597

Db	1442	CGGGGGTGGGAAACGACAGAGGGAGGGGTCCCAAGTCTCACTCTGTAGACAGAGGCC	1501
QY	1598	-----CAGACGACATCGGCTGCA	1615
Db	1502	ACAGAGGTACAAAGACAAACCCAGGCCCTTCTCCCTCTCTACACAGACAGCATCGGCTGCA	1561
QY	1616	CCCTTTGTGGCCAGAGATGAGTGGTCCCGAGCGAAGCACACGCTGCTTCCGCGCAGAGT	1675
Db	1562	CCCTTTGTGGCCAGAGATGAGTGGTCCCGAGCGAAGCACACGCTGCTTCCGCGCAGAGT	1621
QY	1676	CTGGGTTCCTGGGCAATGGGGGAGACCGGAGTGTGTGTGTCTCGGCGGCGAGACCGG	1735
Db	1622	CTCGGTTCCTGGGCAATGGGGGAGACCGGAGTGTGTGTGTGTCTCGGCGGCGAGACCGG	1681
QY	1736	CGCGGGGCTTGTGCTGGCTGGCTTGTGGGGCTGTTCCTTACCATCGGAGCCAGCTAGG	1795
Db	1682	CGCGGGGCTTGTGCTGGCTGGCTTGTGGGGCTGTTCCTTACCATCGGAGCCAGCTAGG	1741
QY	1796	TTTCAAGCCTCGGGGGGGGCCCTGGCCTTGTGGCTGTGGCTGTGGCTGTGGGCTGTGGCTCC	1855
Db	1742	TTTCAAGCCTCGGGGGGGGCCCTGGCCTTGTGGCTGTGGCTGTGGCTGTGGGCTGTGGCTCC	1801
QY	1856	TTCAAGCTCTCTGTTCCTTCCCTGGGACAGCCAGCCCTCCGAGTGCCTGGGCGCCAGAGCCTT	1915
Db	1802	TTCAAGCTCTCTGTTCCTTCCCTGGGACAGCCAGCCCTCCGAGTGCCTGGGCGCCAGAGCCTT	1861
QY	1916	TGTCCACACTCCGCTACAGGGGCTGCTAGACACACTCTTCTCGAGGCGGCCAGAGATCT	1975
Db	1862	TGTCCACACTCCGCTACAGGGGCTGCTAGACACACTCTTCTCGAGGCGGCCAGAGATCT	1921
QY	1976	TCTGTGAGTACAGAACTCCCTGTAGTGGGAGAACCGGCTGAGTGGTCCGCGGGGGG	2035
Db	1922	TCTGTGAGTACAGAACTCCCTGTAGTGGGAGAACCGGCTGAGTGGTCCGCGGGGGG	1981
QY	2036	CGTGGGCTTGAGCTGTGTGCTCTGTGGCATGCTGTGTGAGGTTCGCACTGTGACACTGTT	2095
Db	1982	CGTGGGCTTGAGCTGTGTGCTCTGTGGCATGCTGTGTGAGGTTCGCACTGTGACACTGTT	2041
QY	2096	ACCTGTGTGCTTCCCGCCGGAGGTGTGACAGGACTGTGACATGCTGCCACGAGAGGCG	2155
Db	2042	ACCTGTGTGCTTCCCGCCGGAGGTGTGACAGGACTGTGACATGCTGCCACGAGAGGCG	2101
QY	2156	TGTGTGCACTGGCCGACACAGGTCCTGGGATGAGTTCGGCTACGAGCGACGACCAATGTGCA	2215
Db	2102	TGTGTGCACTGGCCGACACAGGTCCTGGGATGAGTTCGGCTACGCTTCGCGCTTACGCGACGACCAATGTGCA	2161
QY	2216	CGCTGGCCTTCTCTGCTTCTGGGCACTTCTCTGGTGGAGGACAGCCGGGCTGCTACA	2275
Db	2162	CGCTGGCCTTCTCTGCTTCTGGGCACTTCTCTGGTGGAGGACAGCCGGGCTGCTACA	2221
QY	2276	ACCGTGGCCGAGGCTACACTTTTGGCATGCTGGGCTACTTTCACACCTGGGCTCTTGG	2335
Db	2222	ACCGTGGCCGAGGCTACACTTTTGGCATGCTGGGCTACTTTCACACCTGGGCTCTTGG	2281
QY	2336	TGCCCCCTTGTGGCCAAATGTACAGGTGGTCTTACAGGCCGCCGCTGAGATGGCGGCTCC	2395
Db	2282	TGCCCCCTTGTGGCCAAATGTACAGGTGGTCTTACAGGCCGCCGCTGAGATGGCGGCTCC	2341
QY	2396	TGCTCTGTGTCTGGGCACTCTGGCTGTCTTCCACTGCCCCAGGTGTACCTGTCTCATGC	2455
Db	2342	TGCTCTGTGTCTGGGCACTCTGGCTGTCTTCCACTGCCCCAGGTGTACCTGTCTCATGC	2401
QY	2456	GGCAGCCAGGGCTCAACACCCCGAGTCTTCTCTGGAGAGGGGCCCTGGGAGTGGCCAG	2515
Db	2402	GGCAGCCAGGGCTCAACACCCCGAGTCTTCTCTGGAGAGGGGCCCTGGGAGTGGCCAG	2461
QY	2516	GCCAGAAATGACGGGACACAGGAATTCAGGGGAAACATGATGA	2559
Db	2462	GCCAGAAATGACGGGACACAGGAATTCAGGGGAAACATGATGA	2505

RESULT 11

QY 301 GATACGCTGGAGCCTGTGTGGCCATGAGCCCAACCTCATGTTCTGGCCAGGCA 360
DB 301 GACACGACGCTGAGCAGGTGTGACCATGAGCCCAAGCCCATGTTCAATGTGGCAAGGTG 360
QY 361 GGCAGCGCGACATGGCGGCTACTGCAATACAGGAGTACAGCCCGGCTGTGGCT 420
DB 361 GGAAGTCAAGACATGGCTGCTACTGCAACACAGTACCAACCCCGGTGTGGCT 420
QY 421 GTCATGGGCCCCACCTGTGTAGAGCTGCGCATGCTACCGGCAAGTTCTTACGTTCTTC 480
DB 421 GTCATGTGGTCCCACTCATGAGAGCTTGCCTCATTTACAGGCAAGTTCTTACGTTCTTC 480
QY 481 CTGATGCCCCAGGTGACGTAGGTGTGATGATGAGTGTGAGCGCCCGGAGACCTTC 540
DB 481 CTGATGCCCCAGGTGACGTAGGTGTGATGATGAGTGTGAGCGCGGAAACATTT 540
QY 541 CCCCTCTTCTTCCGACCGCTGCCAGGACCTGTGACGTGAGCGGCGGCGGAGAGTG 600
DB 541 CCATCTCTTCTTCCGACAGTCCAGGTGACCGGCTGACGTGACGCGCGTGTGACACTG 600
QY 601 CTGAGAGATTCGGCTGGAACCTGGGTGGCCGCTGGGACAGCAGAGATGACGCGCG 660
DB 601 TTGCAAGATTTTCACTGGAACCTGGGTGGCTGCTTAGGTAGTGAATGATGATGGCGCG 660
QY 661 CAGGCGCTGAGCATCTTCTGCGCCCTGGCGCGGACGCGCATCTGCTATGCGCAGAG 720
DB 661 GAAAGTCTGAGCATCTTCTGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 GGCCTGTGCGCGCTGCGCGCGCGGATGACTGCGGCGTGGGAGAGTGCAGAGCTGCTG 780
DB 721 GGCCTGTGCGCGCTGCGCGCGCGGATGACTGCGGCGTGGGAGAGTGTGATGTGCTA 780
QY 781 CACCAAGTGAACAGAGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 781 CGCCAAAGTGAACAGAGAGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 841 CACCCCTCTTCACTACAGATGACAGAGCTGCTGCGCCAGAGTGTGGGTGGCGAGC 900
DB 841 TACCTCTCTTCACTACAGATGACAGAGCTGCTGCGCCAGAGTGTGGGTGGCGAGT 900
QY 901 GAGGCTGTGCTGACCTGTGACCTGTGATGAGGCTGCGCGGCTGAGGCGCAGATGGGCG 960
DB 901 GAGGCTGTGCTGACCTGTGACCTGTGATGAGGCTGCGCGGCTGAGGCGCAGATGGGCG 960
QY 961 GTGCTGTGCTGCTGCTGACAGAGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 961 GTGCTGTGCTGCTGCTGACAGAGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1021 CACCTGTGCGCGGCGACAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
DB 1021 CCGCTGTGCGCGGCGACAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1081 CTGAGAGAGAGAGCTGTGTGGGCGACGCTGCGCGAGTGTGATGATGACAGCTGCTGAG 1140
DB 1081 CTGAGAGAGAGAGCTGTGTGGGCGACGCTGCGCGAGTGTGATGATGACAGCTGCTGAG 1140
QY 1141 GTGAGCGCAGAGGCTA-----ATCACACAGAGCTTCTCT 1176
DB 1141 GTGAGCGCAGAGGCTA-----ATCACACAGAGCTTCTCT 1176
QY 1177 GTCTACGAGCTGTGTATGAGTGTGGGCGGCGGCTGCGACACATCTTGTGAGAGCGC 1236
DB 1177 GTCTACGAGCTGTGTATGAGTGTGGGCGGCGGCTGCGACACATCTTGTGAGAGCGC 1236
QY 1237 TCAGGCTGTGCGCGGCGAGAGCGGCTGAGAGCTGAGAGCTGCTGAGAGCATGTACAA 1296
DB 1237 TCAGGCTGTGCGCGGCGAGAGCGGCTGAGAGCTGAGAGCTGCTGAGAGCATGTACAA 1296
QY 1297 CTGACCTTCCAGCTGTGCGGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1356
DB 1297 CTGACCTTCCAGCTGTGCGGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1356
QY 1318 ATGAGTTTCCGCTGCTGAGAGCTTGTGATGATGATGATGATGATGATGATGATGATG 1377
DB 1318 ATGAGTTTCCGCTGCTGAGAGCTTGTGATGATGATGATGATGATGATGATGATGATG 1377
QY 1357 GAGTACGACCTGAGAGCTGTGGGTGTGGCAGGCTGACGTGCCAGGCTCCAGAGAGCTGGC 1416

DB 1378 GAATATACCTGGAAGATGTGGGTGTGGCAGAGCCCTACCTGTATACATGATGTAGGC 1437
QY 1417 AGTTTCAACGGGACCTCAGAGAGAGGCGCTGAAGATCCGCTGGCAGAGCTGTGACAA 1476
DB 1417 AGTTTCAACGGGACCTCAGAGAGAGGCGCTGAAGATCCGCTGGCAGAGCTGTGACAA 1476
QY 1438 ACCTTCAAGGACCTTCACTGACCTGACGACCTGCAAAATGATTTGGC-----CAGGCAAC 1491
DB 1438 ACCTTCAAGGACCTTCACTGACCTGACGACCTGCAAAATGATTTGGC-----CAGGCAAC 1491
QY 1477 CAGAAGCCGCTGTCCGCGTGTGCGGAGTGTGCGAGAGGCGGAGGTGTGGCGGCGTCAAG 1536
DB 1477 CAGAAGCCGCTGTCCGCGTGTGCGGAGTGTGCGAGAGGCGGAGGTGTGGCGGCGTCAAG 1536
QY 1492 CAGGTGCGAGTGTCCGAGTGTGTCCGCGAGTGTGCAAAAGATGTGGCGAGGTGAGAGTAA 1551
DB 1492 CAGGTGCGAGTGTCCGAGTGTGTCCGCGAGTGTGCAAAAGATGTGGCGAGGTGAGAGTAA 1551
QY 1537 GGGTTCACACTCTCTGCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1596
DB 1537 GGGTTCACACTCTCTGCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1596
QY 1552 GGGTTCATTCCTCTGCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1611
DB 1552 GGGTTCATTCCTCTGCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1611
QY 1597 CCAGAGCAGATGCGCTGACCTTTGTGGCCAGAGATGATGTGTGTGTGTGTGTGTGTGTGT 1656
DB 1597 CCAGAGCAGATGCGCTGACCTTTGTGGCCAGAGATGATGTGTGTGTGTGTGTGTGTGTGT 1656
QY 1612 CCAGATGACTTCACTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1671
DB 1612 CCAGATGACTTCACTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1671
QY 1657 CGCTGTTCGCGCGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1716
DB 1657 CGCTGTTCGCGCGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1716
QY 1672 ACCTGTACCTGCGAGGCGCAAGTTTGTGGCTGTGGGCGGAGCGAGCTGTGCTGCTGCT 1731
DB 1672 ACCTGTACCTGCGAGGCGCAAGTTTGTGGCTGTGGGCGGAGCGAGCTGTGCTGCTGCT 1731
QY 1717 CTCTGTGCTGCTGAGCTGTGGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 1776
DB 1717 CTCTGTGCTGCTGAGCTGTGGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 1776
QY 1732 CTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1791
DB 1732 CTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1791
QY 1777 CATGAGGACAGCCAGCTGTGTGAGGCTGTGGGCGGCGGCGGCTGCTGCTGCTGCTGCTG 1836
DB 1777 CATGAGGACAGCCAGCTGTGTGAGGCTGTGGGCGGCGGCGGCTGCTGCTGCTGCTGCTG 1836
QY 1792 TACGTGGAGAGCCCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1851
DB 1792 TACGTGGAGAGCCCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1851
QY 1837 TGCTGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1896
DB 1837 TGCTGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1896
QY 1852 TGCTGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1911
DB 1852 TGCTGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1911
QY 1897 TGCTGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1956
DB 1897 TGCTGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1956
QY 1912 TGCTGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1971
DB 1912 TGCTGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1971
QY 1957 CTGAGGCGGCGGAGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2016
DB 1957 CTGAGGCGGCGGAGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2016
QY 1972 CTGAGAGCGCGAGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2031
DB 1972 CTGAGAGCGCGAGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2031
QY 2017 AGTGTGCTGCTGCGGCGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 2076
DB 2017 AGTGTGCTGCTGCGGCGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 2076
QY 2032 TGACGACTACCTTTCGCGGCGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 2091
DB 2032 TGACGACTACCTTTCGCGGCGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGT 2091
QY 2077 GTGCGACGTGTGACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2136
DB 2077 GTGCGACGTGTGACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2136
QY 2092 GCTGACCTATGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2151
DB 2092 GCTGACCTATGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2151
QY 2137 ATGCTGCGCAGAGAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2196
DB 2137 ATGCTGCGCAGAGAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2196
QY 2152 GTGCTGCGCAGAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2211
DB 2152 GTGCTGCGCAGAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2211
QY 2197 GCGCAGCGCAGATGCGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2256
DB 2197 GCGCAGCGCAGATGCGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2256
QY 2212 GTGCAATGACCAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2271
DB 2212 GTGCAATGACCAATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2271
QY 2257 AGCGAGCGGCGCTGTGACCAACGCTGTGCGGCGGCTGTGACCTTGTGACATGCGGCTG 2316
DB 2257 AGCGAGCGGCGCTGTGACCAACGCTGTGCGGCGGCTGTGACCTTGTGACATGCGGCTG 2316
QY 2272 AGCGAGCGGCTGTGACCAACGCTGTGCGGCGGCTGTGACCTTGTGACATGCGGCTG 2331
DB 2272 AGCGAGCGGCTGTGACCAACGCTGTGCGGCGGCTGTGACCTTGTGACATGCGGCTG 2331
QY 2317 ATCACGCTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2376
DB 2317 ATCACGCTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2376
QY 2332 ATCATGTGGGCTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2391
DB 2332 ATCATGTGGGCTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2391
QY 2377 GTGCAATGTGGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2436
DB 2377 GTGCAATGTGGGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2436
QY 2392 GTGCAATGTGGGCTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2451
DB 2392 GTGCAATGTGGGCTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2451
QY 2437 AGTGTATACCTGCTCATGCGGCGGAGGCTGCAACGCCCGGAGTCTTCTGCTGGAGGG 2496
DB 2437 AGTGTATACCTGCTCATGCGGCGGAGGCTGCAACGCCCGGAGTCTTCTGCTGGAGGG 2496

Db 2452 AATGCTATGACTTCTGTGGTGGCCAGAGCTACACACCCAGAGATGCTCTGGAGAG 2511
 Oy 2497 GGGCCGTGGGATGCCCAAGGCCAGAAATGACGGGA 2530
 Db 2512 AGCCCAAGAGACATCAAGATGGGATAGTGGA 2545
 RESULT 13
 AAS97393
 ID AAS97393 standard; cDNA: 2577 BP.
 AC AAS97393;
 DT 12-MAR-2002 (first entry)
 DE Murine SACL gene cDNA.
 KW Human; mouse; SACL; carbohydrate; sweetener; ethanol; alcoholism; ss;
 obesity; diabetes; transgenic embryo; body tissue; body fluid; pancreas;
 blood; tongue; PCR primer; anorectic; antidiabetic; gene therapy;
 protein replacement therapy.
 OS Mus sp.
 PN W020183749-A2.
 PD 08-NOV-2001.
 PE 25-APR-2001; 2001WO-US13387.
 PR 28-APR-2000; 2000US-200794P.
 PR 28-JUL-2000; 2000US-221419P.
 PR 10-NOV-2000; 2000US-247443P.
 PA (WARNER) WARNER LAMBERT CO.
 PA (NONE-) MONELL CHEM SENSES CENT.
 PI Bachmanov AA, Beauchamp GK, Chatterjee A, De Jong PJ, Li S, Li X;
 PI Ohnen JD, Reed DR, Ross D, Tordoff MG;
 DR WPI: 2002-075162/10.
 DR P-PSDB: AAU73183.
 PT Novel isolated polypeptide comprising variant form of mouse or human
 PT SACL polypeptide, and is associated with altered preference for
 PT carbohydrates or other sweeteners, useful for preventing obesity,
 PT diabetes, alcoholism -
 PA Claim 1; Page 28-29; 239pp; English.
 PS The invention relates to an isolated polypeptide, comprising a variant
 PS form of mouse or human SACL polypeptide. The variant form is associated
 PS with altered preference for carbohydrates, other sweeteners or ethanol.
 PS The polypeptide and its associated DNA sequence can be produced by
 PS recombinant techniques and is useful for preventing obesity, diabetes or
 PS alcoholism associated with SACL expression. The sequences are useful in
 PS screening for drugs and sweeteners. Recombinant cell lines and transgenic
 PS embryos may be used in screening for and identifying agents that induce
 PS or repress function of SACL. Predisposition to diabetes, obesity or
 PS alcoholism can be ascertained by testing any fluid or tissue of a human
 PS (such as blood, pancreas or tongue) for sequence variations of the SACL
 PS gene. A sequence variation of the SACL locus may indicate a
 PS predisposition to diabetes, obesity and/or alcoholism and may provide a
 PS diagnostic mark. The polynucleotide can be detected in a biological
 PS sample by contacting the DNA with a probe to form a hybridisation complex
 PS which is then detected. The sequences represent cDNA encoding human and
 PS mouse SACL polypeptides and PCR primers specific for the SACL genes.
 SQ Sequence 2577 BP; 519 A; 735 C; 689 G; 634 T; 0 other;

Oy 1 ATGCTGGGCGCTGCTGTGGGCTCAAGCCCTGAGGCTCTGACCTGGAGCGGG 60
 Db 1 ATGCCAGCTTTGGCTATGATGAGTGTCTCAGCCGTCTTCTGAGCTGGAGTGGGG 60
 Oy 61 GCGCCATTTGGCTGTCTCAGCACTAGATGAGAGGGGAGATACGTCTGGGGGGCTG 120
 Db 61 GCGCTTTGTGTCTCAGCACTAGATGAGAGGGGAGATACATATGCGGGGCTA 120
 Oy 121 TTCCCTTGGGGAGGCGGAGAGGCTGCGCTCGCAGCCGAGACAGGCCAGCCCT 180
 Db 121 TTTCCCTGGGCTCAACGAGAGGCGCAGCTCAACAGAGAACACCAACCAAGCATC 180
 Oy 181 GTGTGACAGGATGCTCTCAACAGGCTGCTGTGGGACAGTCCATTAATATGGCCGTG 240
 Db 181 CCGTCAACAGGATGCTCAACCCCTTGGTGTCTGGGCACTGCTATTAAGATGGCTGTG 240
 Oy 241 GAGGAGATCAACCAACCACTGATGCTGCGCGGGCTGCGGCTGGCTAGCACTCTTT 300
 Db 241 GAGGAGATCAACCAACCACTGATGCTGCGCGGGCTGCGGCTGGCTAGCACTATTT 300
 Oy 301 GATACGCTCTGGAGCCTGTGTGGCCATGAAGCCAGCTCATTTCTTGGCCAGGCA 360
 Db 301 GACACATGCTCGAGCCAGGCTGACATGAATTCAGTCTCATGTTCTGGCCAGGCTG 360
 Oy 361 GGCAGCGCGGACATGCGCGGCTGACATGACAGGACGACAGCCCGGTGCTGGCT 420
 Db 361 GGCAGTCAACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 Oy 421 GTCATGCGGCGCCACTGCTGACAGCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCT 480
 Db 421 GTCATGCGGCGCCACTGCTGACAGCTGCGCATGCTGCTGCTGCTGCTGCTGCTGCT 480
 Oy 481 CTCATGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Db 481 CTCATGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 Oy 541 CCCCTCTCTTCCGACCGCTGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 600
 Db 541 CCATCTCTTCCGACCGCTGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 600
 Oy 601 CTGACAGAGTTCGCTGAGTGGTGGCGCGCTGCGGCGGCTGCGGCGGCTGCGGCGG 660
 Db 601 TTGACAGACTTCACTGAGTGGTGGCGCGCTGCGGCGGCTGCGGCGGCTGCGGCGG 660
 Oy 661 CAGGCGCTGAGCATTTCTGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGGCTGCGG 720
 Db 661 GAGGCTGAGCATTTCTGAGTGGTGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGG 720
 Oy 721 GAGGCTGAGCATTTCTGAGTGGTGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGG 780
 Db 721 GAGGCTGAGCATTTCTGAGTGGTGGCGGCTGCGGCGGCTGCGGCGGCTGCGGCGG 780
 Oy 781 CACGAGTGAACCAAGAGAGGCTGAGTGGTGGCTGAGTGGTGGCTGAGTGGTGGCTG 840
 Db 781 CCGCAAGTGAACCAAGAGAGGCTGAGTGGTGGCTGAGTGGTGGCTGAGTGGTGGCTG 840
 Oy 841 CAGGCGCTTCACTAAGATGATGACAGAGCTTGGCCCAAGTGGTGGTGGCTGAGC 900
 Db 841 TACGCTCTTTTAAATTAAGATGATGACAGAGCTTGGCCCAAGTGGTGGTGGCTGAGC 900
 Oy 901 GAGGCGCTTCACTTGAAGTGGTGGCTGAGTGGTGGCTGAGTGGTGGCTGAGTGGCTG 960
 Db 901 GAGTGGTGGCTTGAAGTGGTGGCTGAGTGGTGGCTGAGTGGTGGCTGAGTGGCTG 960
 Oy 961 GTGCTTGGCTTCTCAGAGAGGCTGCGGCTGAGTGGTGGCTGAGTGGTGGCTGAGTGG 1020
 Db 961 GTGCTTGGCTTCTCAGAGAGGCTGCGGCTGAGTGGTGGCTGAGTGGTGGCTGAGTGG 1020
 Oy 1021 CACCTGGGCGCTGCGGCGGCTGAGTGGTGGCTGAGTGGTGGCTGAGTGGTGGCTGAG 1080
 Db 1021 CACCTGGGCGCTGCGGCGGCTGAGTGGTGGCTGAGTGGTGGCTGAGTGGTGGCTGAG 1080

Query Match 56.9%; Score 1454.8; DB 24; Length 2577;
 Best Local Similarity 74.3%; Pred. No. 5.2e-238;
 Matches 1891; Conservative 0; Mismatches 622; Indels 33; Gaps 3;

QY 1081 CTGAGGAGGAGCTGTGGGCGCAGCTGCCGAGTGTGACTGATCAGC----- 1132
 DB 1078 CTGAGGAGGAGATGATGATGGGCAACGCTGTCTCAAGGTGTGACGATCATGCTCAGAAC 1137
 QY 1133 -----TGCAAGAGTGAAGGCGAGGGCTAAATCAACCAACGAGCTTCTCT 1176
 DB 1138 CTATCATCTGGGCTGTGTGAGAACCTATCAGCTGGGCAATTTGCAACCAAAATTTTGA 1197
 QY 1177 GTCTACGACCTGTGTATACGCTGTGGCCAGGCGCTGCACAACACTCTTTCAGAACGCC 1236
 DB 1198 ACCTATGACACTGTGTACAGTGTGGGCTCAGGCGCTTCAACAACCTCAGTGAATGTCTC 1257
 QY 1237 TCAGGCTGCCCGCGCAGAACCCCGTGAAGCCCTGGCAGCTCTGTGAAGCATGTACAC 1286
 DB 1258 TCACATTTGCCACGTATCAAGAACATGTTCTACCCCTGGCAGCTCCGGAACATGTACAT 1317
 QY 1297 CTGACCTTCACGCTGGGCGGCGCTGCGTGGTGGACAGCAGCGGAAGAGTGCATG 1356
 DB 1318 ATGAGTTTCATGCTCAGACTTGACACTACAGTTGATGCTGAAGGGAATGTAGCATG 1377
 QY 1357 GAGTACGACCTGAAGCTGTGGGTGTGGCAGGGCTCAGTGCACAGGCTCCACAGAGTGGC 1416
 DB 1378 GAAATATACCTGAAGATGTGGGTGTGGCAGAGCCCTACACTGTATTTACATAGTGTGGC 1437
 QY 1417 AGGTTCAAGGCGCAGCTCAGAGACAGAGCGCTGAAGATCCGCTGGCAGACGTTGACAC 1476
 DB 1438 ACCTTCAAGCGCACCCCTTACAGCTCAGCAGTCTAAATGTACTGGC-----CAGGCAAC 1491
 QY 1477 CAGAAAGCCGTCGCCGCTGTGCGGCGAGTGCAGAGAGGCGGAGTGCCTGGGCTCAAG 1536
 DB 1492 CAGGTGCGACCTTCCAGTGTCCCGCAGTGCAGAAAGATGGCCAGTTCGCGAGTAAAG 1551
 QY 1537 GGGTTCCACTCTGCTCTAGACTGTGTGAGACTGCAGAGCGGCGACACTACCGGCAAAAC 1596
 DB 1552 GGGTTTCATTCTGCTCTATGACTGTGGTGAAGTGGTGAAGGCGGCGACACTACCGGAACAT 1611
 QY 1597 CCAGACGACATCGCTCACCCTTTTGGCGCAGATAGTGTGCTCCCGGAGCGACAGACA 1656
 DB 1612 CCAGATATCACTTCACTTACTTCACTGTATACAGAGACAGTGTGCCAGAAAGACACA 1671
 QY 1657 CGGTGCTTCCGCGCAGAGTGTGCGTCTGTGCATGAGGCGGCGGCTGTGCTCTGTG 1716
 DB 1672 GCGTGTCTTACTCGCAGAGCCCAAGTTTCTGTGGGCGGAGCAGTGTGCTCTGCTC 1731
 QY 1717 CTCTCTGCTGTGAGCTGTGGGCTGTGGGCTGTGTGCTGTGCTGTGCTGTGCTGTGCT 1776
 DB 1732 CTCTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1791
 QY 1777 CATTGGGAGACCCACATGCTTCAAGGCTCGGGGGGGGCGCTCGGCTCTTGTGGCTGTG 1836
 DB 1792 CACTGGGAGACGCCCTTGTTCAGGCGCTCAAGTGTGCTCAGATTTCTGTGGCTGTATC 1851
 QY 1837 TGCTTGGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1896
 DB 1852 TGCTTGGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1911
 QY 1897 TGCTTGGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1956
 DB 1912 TGCTTGGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1971
 QY 1957 CTGAGGCGGCGCAGATCTTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2016
 DB 1972 CTGAGGCGGCGCAGATCTTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2031
 QY 2017 AGTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2076
 DB 2032 TGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2091
 QY 2077 GTGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2136
 DB 2092 GCAGCACTATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 2151
 QY 2137 ATGCTGCCACGAGAGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 2196

DB 2152 GTGCTGCCACAGAGTACTGAGACACTGCGAGTGTCTGTGGGTGCTGAGCTGTGGCTTG 2211
 QY 2197 GCGAGCGCACCAATGCGACAGCGTGGCTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 2256
 DB 2212 GTGCACTACCAATGCGACATGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2271
 QY 2257 AGCGAGCGGCGCTGTCAACACCGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2316
 DB 2272 AGCGAGCGTGGCGCTGTCAACACCGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 2331
 QY 2317 ATCACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 2376
 DB 2332 ATCACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 2391
 QY 2377 GTGCAATGGGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 2436
 DB 2392 GTGCAATGGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 2451
 QY 2437 AGGTGTACCTGTCTCATGCGGCAAGCCAGGCTCAACACCCCGAGTCTTCTGTGGAGGG 2496
 DB 2452 AAGTGTATGTGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 2511
 QY 2497 GCGCTGTGGATGCCCAAGGCCAGAA 2522
 DB 2512 AATGCCAAGAAAGCAGCAGATGAGAA 2537
 RESULT 14
 AAD34458
 ID AAD34458 standard; cDNA; 1113 BP.
 XX
 AC AAD34458;
 XX
 DT 16-JUL-2002 (first entry)
 XX
 DE Human GPCR cDNA #1.
 XX
 KW Human; G protein-coupled receptor; GPCR; central nervous system disease;
 KW cancer; anorexia; bulimia; asthma; benign prostatic hypertrophy; ulcer;
 KW allergy; acute heart failure; myocardial infarction; urinary retention;
 KW Parkinson's disease; Alzheimer's disease; hypotension; angina pectoris;
 KW hypertension; osteoporosis; inflammation; infection; dementia; diabetes;
 KW Leukemia; lymphoma; gene therapy; gene; ss.
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT 1.1113
 FT CDS /tag- a
 FT /product- "Human GPCR protein #1"
 PD
 XX
 PD W0200224885-A2.
 XX
 PD 28-MAR-2002.
 XX
 PF 21-SEP-2001; 2001MO-EP10929.
 XX
 PR 25-SEP-2000; 2000US-234982P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Kossida S;
 XX
 DR WPI: 2002-339949/37.
 DR P-PSDB: AAE21817.
 DR
 PT New G-protein coupled receptor polynucleotide and its encoded protein,
 PT useful for identifying modulators of GPCR activity, and in gene therapy
 PT for treating bacterial infection, cancer, acute heart failure or
 PT inflammation -
 XX
 PS Claim 19; Fig 1; 101bp; English.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 12:30:29 ; Search time 3569.38 Seconds

(without alignments)
11611.043 Million cell updates/sec

Title: US-09-927-315-14

Perfect score: 2559

Sequence: 1 atgctggccctgctgtctc.....atcaggggaacatgagta 2559

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1615406 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*

1: em_estba:**
2: em_esthum:**
3: em_estnu:**
4: em_estnu:**
5: em_estnu:**
6: em_estnu:**
7: em_estnu:**
8: em_estnu:**
9: gb_est1:**
10: gb_est2:**
11: gb_est3:**
12: gb_est3:**
13: gb_est4:**
14: gb_est5:**
15: em_estfun:**
16: em_estom:**
17: gb_gss:**
18: em_gss_hum:**
19: em_gss_inv:**
20: em_gss_pln:**
21: em_gss_vit:**
22: em_gss_vit:**
23: em_gss_vit:**
24: em_gss_mam:**
25: em_gss_mus:**
26: em_gss_other:**
27: em_gss_pro:**
27: em_gss_rnd:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	367.8	14.4	750 14	B0178590 UI-M-EVO-
2	348.4	13.6	558 12	BE754542 BE754542
3	219.2	8.6	323 10	AM417577 AM417577
4	108	4.2	973 17	CNS02BDU CNS02BDU
5	107.4	4.2	938 13	B1462667 B1462667
6	106.6	4.2	612 13	BM439858 BM439858

7	106	4.1	495	14	BM706305
8	104.8	4.1	298	9	AT562167
9	104.6	4.1	709	12	BG182090
10	102.2	4.0	168	9	AA907022
11	101.8	4.0	1038	17	CNS05068
12	99.2	3.9	464	9	AA937218
13	99.2	3.9	498	10	AM051287
14	98.6	3.9	635	10	AM771483
15	97.2	3.8	738	12	BG191110
16	96.8	3.8	694	12	BG194102
17	86	3.4	830	17	AZ750699
18	78.6	3.1	1041	17	CNS012X1
19	75.6	3.0	829	17	CNS04565
20	74	2.9	718	17	AZ972907
21	73	2.9	902	17	CNS04KSY
22	71.4	2.8	573	9	AT657897
23	69.2	2.7	525	17	AZ483105
24	69	2.7	1470	12	BE910284
25	68.4	2.7	602	9	AT434785
26	67	2.6	562	9	AT742401
27	66.6	2.6	925	17	CNS0091P
28	66.6	2.6	1014	17	CNS033XY
29	65	2.5	487	13	BM259415
30	64.6	2.5	479	13	BM259601
31	63	2.5	935	17	CNS006XK
32	62.6	2.4	571	10	AM453329
33	62.6	2.4	1244	12	BG846745
34	62.4	2.4	1085	17	AG152796
35	62.2	2.4	1034	11	AY111284
36	62	2.4	820	10	BE558912
37	61.2	2.4	545	12	BG414601
38	61.2	2.4	656	13	B1874236
39	61.2	2.4	700	9	AL507686
40	61.2	2.4	810	12	BF628952
41	61.2	2.4	925	17	CNS0091P
42	61	2.4	635	14	BQ768749
43	60.8	2.4	888	17	CNS03200
44	60.6	2.4	454	17	AQ466269
45	60.4	2.4	657	14	BQ245583

ALIGNMENTS

RESULT 1
B0178590/c 750 bp mRNA linear EST 30-APR-2002
UI-M-EVO-bws-h-23-0-UI.r1 NIH.BMAP_EVO Mus musculus CDNA clone
IMAGE:5701318 5', mRNA sequence.

ACCESSION B0178590
VERSION B0178590.1 GI:20354082

KEYWORDS EST.
SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 750)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scuriongnathi; Muridae; Murinae; Mus.

NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa

CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.
Location/Qualifiers

FEATURES

[illegible]

DB	70	GAACATGTCAGTGGGCAACGCTGTCACAGGTCGTACAGACATCATCTCGAAGAACTATCA	11
QY	1147	GCAGGGCT 1154	
DB	10	TCTGGGCT 3	
RESULT 2			
LOCUS	BE754542	558 bp	linear
DEFINITION	208082 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.		EST 25-APR-2001
ACCESSION	BE754542		
VERSION	BE754542.1	GI:10168534	
KEYWORDS	EST.		
SOURCE	COV.		
ORGANISM	Bos taurus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.		
AUTHORS	Smith,T.P.L., Grosse,W.M., Fekling,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.M., Rohrer,G.A., Chitto-McKown,C.G., Petrea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.		
TITLE	Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle		
JOURNAL	Genome Res. 11 (4), 626-630 (2001)		
MEDLINE	21180013		
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -mismore 18 and -mismatch 12 options. PCR Primers FORWARD: AGGAACAGCTATGACCAT BACKWARD: GTTTCACAGTCACGAG Plate: 54 row: A column: 22 Seq primer: ATTATGTCGACATATAC. Location/Qualifiers 1..558 /organism="Bos taurus" /db_xref="taxon:9913" /clone_lib="MARC 2BOV" /tissue_type="pooled" /lab_host="DH10B" /note="Vector: pCMV SPORT6, site_1: NotI, site_2: SalI; library made from pooled tissue from testis, thymus, semitendinosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."		
FEATURES	source		
BASE COUNT	98 a	177 c	179 g 104 t
ORIGIN			
Query Match	13.6%;	Score 348.4;	DB 12; Length 558;
Best Local Similarity	76.5%;	Pred. No. 2.1e-52;	
Matches 427;	Conservative 0;	Mismatches 131;	Indels 0; Gaps 0;
QY	1277	TCTCTGAGAACATGTACAACTGACCTTCACAGTGGGCGGCTGCGGCTGCGGTTGACAA	1336
DB	1	TCTCTGAGAACATGTACAACTGAGTTTCACAGCATACAAACCGGACCTGCGAGTTGACAA	60
QY	1337	GCAGCGGAAGCTGGACATGAGTACGACCTGAGCTGTGGTGCGACGGGCTTAGTGC	1396
DB	61	CCAGCGGAAGTGTGACCTTGGCTTGTGACCTGACCTGTGGGTTTGAGAGGGACCGGATGC	120
QY	1397	CCAGGCTCCAGAGCTGGGCGAGCTCAACGGAGGCTCAGAGCAAGAGGGCCGGAAGATCC	1456
DB	121	CCATGCTCGGACCTGTGGGCACTTCAACGGGAGCGAGCTGAGAGCTCCAGTTTCCACACATGA	180

QY 1457 GCTGGACACGCTGTGACACACAGAACCCCTGTCCCGGTCTCGCGGAGTGCAGAGG 1516
 DB 181 TGTGGACACACCCAGGAAACACAGAGCCCGTCTCCAGTGTCTACAGGAGTGCAGAGG 240
 QY 1517 GCCAGGTGCGCGGGCTCAAGGGGTTCCACTCTGTCTGTACGACTGTGTGACTCGAG 1576
 DB 241 GCCAGGTGCGCGGGCTCAAGGGGTTCCACTCTGTCTGTACGACTGTGTGACTCGAG 300
 QY 1577 GGGGAGCTACGGGGAACCAAGAGACATGCGCTGTCTGTGTGCGCGAGGAGT 1636
 DB 301 CAGGAGCTACGAGGAGGACACAGAGATGCCCTGTGACAGAGTGTGACAGGAGT 360
 QY 1637 GGTCCCGGAGGACACACACGCTGTCTCCGCGAGTGTGTCTGTCTGTGAGTGGGCG 1696
 DB 361 GGTCCCGGAGGAGTGTACCGGCTGTCTCCGCGAGGAGTGTGTCTGTGAGTGGGCG 420
 QY 1697 AGCCGCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 1756
 DB 421 AGCCGCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 480
 QY 1757 CTTGGGCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 1816
 DB 481 CCTGGGAGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 540
 QY 1817 TGGCCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 1834
 DB 541 GGGCCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 558

RESULT 3
 AM417577 323 bp mRNA linear EST 09-JUL-2000
 LOCUS 54413 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
 DEFINITION AM417577
 ACCESSION AM417577.1 GI:6945459
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine
 Unpublished (2000)
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smithemall.marc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 20 and -mismatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCAGTACGAGG
 Plate: 20 row: E column: 14
 Seq primer: ATTGAGTGACATATAG.

FEATURES
 source
 1. .323
 Location/Qualifiers
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="MARC 1P1G"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI; Library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos."

BASE COUNT 43 a 120 c 85 g 72 t 3 others
 ORIGIN

Query Match 8.6%; Score 219.2; DB 10; Length 323;
 Best Local Similarity 82.3%; Pred. No. 2.1e-29;
 Matches 251; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
 QY 2127 GCACTGGACATGCTGCCACAGGAGCGCTGTGTGACTGCCACACGCTCTGTGGTGA 2186
 DB 19 GGACTGGGCTGTGCTACCCACANAGCACTGTGTGACTGCCGCGTCACTCTGTGATCAG 78
 QY 2187 CTTGGGCTTACGGGACGCCCAATGACAGCTGTGTCTGTCTGTCTGTCTGTCTGT 2246
 DB 79 CTTGAGCTGTGTGATGTGTGCAATGACAGCTGTGTCTGTCTGTCTGTGTGTGT 138
 QY 2247 CTTGGTGGGAGGAGCGGGGTGTGTACACGCTGTGTGTGTGTGTGTGTGTGTGT 2306
 DB 139 CTTGTGTGACAGCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 198
 QY 2307 GGCCTACTTCATCACCCTGGGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2366
 DB 199 GGCTACTTCATCACCCTGGGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 258
 QY 2367 CAGGCGCCGCTGTGAGATGGGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2426
 DB 239 CCAGCTGTGTGACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 318
 QY 2427 CCACC 2431
 DB 319 CCACC 323

RESULT 4
 CNS02BUU 973 bp DNA linear GSS 12-MAY-2000
 LOCUS 254011 of library G from Tetraodon nigroviridis, genomic survey
 DEFINITION
 ACCESSION AL190335.1 GI:7828439
 VERSION AL190335
 KEYWORDS
 SOURCE
 ORGANISM
 GSS: genome survey sequence.
 Tetraodon nigroviridis.
 Tetraodon nigroviridis.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 1 (bases 1 to 973)
 Roest-Crollius, H., Jallou, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Bernot, A., Fizames, C., Mincker, P., Brottier, P., Quetier, F.,
 Saurin, W. and Weissenbach, J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence
 Unpublished
 2 (bases 1 to 973)
 Roest-Crollius, H., Jallou, O., Dasilva, C., Fizames, C., Fisher, C.,
 Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
 Weissenbach, J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Unpublished
 3 (bases 1 to 973)
 Direct Submission
 Submitted (12-APR-2000)
 Genoscope.

FEATURES
 source
 1. .973
 Location/Qualifiers
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone_lib="254011"
 /clone_lib="G"
 This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/tetraodon.

/note="Genoscope sequence ID : COAG254AH06SP1-end :
PUC-ori"
BASE COUNT 184 a 306 c 247 g 227 t 9 others
ORIGIN

Query Match 4.2%; Score 108; DB 17; Length 973;
Best Local Similarity 50.7%; Pred. No. 1.6e-09;
Matches 359; Conservative 4; Mismatches 329; Indels 16; Gaps 4;

QY 1633 GAGTGTCTCCCGGAGGAGACAGACGCTTCCGCGCCAGGCTCTGCTTCTGAGCATG 1692
DB 280 GAGTGTCTCCCGGAGGAGACAGACGCTTCCGCGCCAGGCTCTGCTTCTGAGCATG 339
QY 1693 GCGCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1752
DB 340 AAGGACACCTTTCGCGGCTTCTGATGCTTCTGATGCTTCTGATGCTTCTGATGCTTCTG 399
QY 1753 GCTGCTTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1812
DB 400 TWGACATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 459
QY 1813 CCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1872
DB 460 CCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 519
QY 1873 CCGTGGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 1932
DB 520 GTGGGCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 579
QY 1933 ACGGGCTGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1992
DB 580 ACTGTGTGTATCTGTGTATCTGTGTATCTGTGTATCTGTGTATCTGTGTATCTGTGTAT 639
QY 1993 CCTGTGAGCTGGGAG 2052
DB 640 AACCCTGGGCTGAG 696
QY 2053 GTGCTGTGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2112
DB 697 GTGCTGTGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 750
QY 2113 CCGGAGGCTGAGGAG 2172
DB 751 CCGTACATCAACACCATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 804
QY 2173 CGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2231
DB 805 AACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 864
QY 2232 CTTCCTGGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2291
DB 865 CTTCCTGGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 924
QY 2292 CACCTTGGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2339
DB 925 CACCTTGGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 972

RESULT 5
LOCUS B1462667 938 bp mRNA linear EST 21-AUG-2001
DEFINITION 603202107F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5268264 5',
ACCESSION B1462667
VERSION B1462667.1 GI:15253323
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 938)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gcrabs-remail.nih.gov

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L16M1676 row: m column: 01
High quality sequence stop: 739.
Location/Qualifiers

FEATURES
SOURCE 1..938
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5268264"
/clone_11b="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescript (modified
pBluescript KS+); Site 1: BamHI, Site 2: SalI; xhoI (gtcgag
); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTVA-3',
size-selected for average insert size 2.2 kb and
normalized to ROI 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 186 a 260 c 294 g 198 t
ORIGIN

Query Match 4.2%; Score 107.4; DB 13; Length 938;
Best Local Similarity 55.9%; Pred. No. 2e-09;
Matches 204; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 367 CGGACATCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 426
DB 147 CACCATATAGAGCTTCAAG 206
QY 427 GGGGCCACTGTGAGAGCTGCGCATGCTGCGGCAAGTCTTACCTTCTGCTGATG 486
DB 207 GGGCCCTGACACCAACCGTGTGCGACACACCGCCGCTGCTGAGCCCTTCTGCTG 266
QY 487 CCGGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 546
DB 267 CCGATGATATACATGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 326
QY 547 TTCTTCCGACCGTCCAGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606
DB 327 TTCTTCCGACCGTCCAGCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 386
QY 607 GAGTGTGGGCTGGAAGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 666
DB 387 AAGTGTGGGCTGGAAGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 446
QY 667 CTGAGCATCTTCTGCGCTTGGCGCGGAGAGCGGAGCTGAGCTGAGCTGAGCTGAGCTGAG 726
DB 447 GTGAGGAGACTGAGAGAACAG 506
QY 727 GTGCC 731
DB 507 ATGCC 511

RESULT 6
LOCUS BM439858 612 bp mRNA linear EST 01-FEB-2002
DEFINITION pgrin.pK001.h19 Normalized Chicken Reproductive Tract cDNA Library
(pgrin) Gallus gallus cDNA clone pgrin.pK001.h19 5' similar to
gd1AD18069.1 (AF127389) putative taste receptor TR1 [Rattus
norvegicus], mRNA sequence.
ACCESSION BM439858

VERSIONS	KEYWORDS
BMA39858.1	GI:18470633
SOURCE	chicken.
ORGANISM	Gallus gallus
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
AUTHORS	1 (bases 1 to 612)
TITLE	Cogburn, L.A. and Nys, Y.
JOURNAL	ESTs from Normalized Chicken Reproductive Tract cDNA library-University of Delaware and INRA, Tours-Poultry Unit Project unpublished (2002)
COMMENT	Contact: Larry A. Cogburn University of Delaware Townsend Hall, Newark, DE 19717, USA Tel: 302-831-1335 Fax: 302-831-2822 Email: cogburnudel.edu, www.chickest.udel.edu.
FEATURES	Location/Qualifiers
Source	1..612
	/organism="Gallus gallus"
	/strain="Commercial broiler and layer"
	/db_xref="taxon:9031"
	/clone="pgrln.pK001.h19"
	/library="Normalized Chicken Reproductive Tract cDNA library (pgrln)"
	/sex="Male and Female"
	/tissue_type="Testis, ovary and oviduct"
	/dev_stage="Various stages; embryonic, post-hatch, immature and sexually-mature"
	/lab_host="E. coli EMEDH10B"
	/note="Vector: pCMVSPORT6; Library made from three total RNA pools from each tissue (testis 25%, ovary 25%, and oviduct 50% of final RNA pool); Single pass sequencing from 5'-end"
BASE COUNT	78 a 241 c 202 g 86 t 5 others
ORIGIN	
Query Match	4.2%; Score 106.6; DB 13; Length 612;
Best Local Similarity	57.5%; Pred. No. 2.6e-09;
Matches 206; Conservative	0; Mismatches 149; Indels 3; Gaps 1;
OY	ACGGCTGCTGTGGGCATCGCCATGAATAAGCCGTCGAGAGATCAACAACAGTGGC 262
DB	ACGGGTACCACCTCTCGCAAAATGAGCCCTTCGCCGTGAGAGATCAACAACCTCACAGC 316
OY	ATCTGCTGCCCCGGGTGGCGCTGAGCCATGACAGCATTTGATACGTCGTCGAGCCTGNG 322
DB	CGCTCTCCCAACGCTACAGCTGCGGTACGAGATCCAGACACTTGACAGGAGGCT--GG 373
OY	TGGCCATGAAGCCCAAGCCCTCATGTCTCTGCGCCAAGGACGAGCCGACATCCGCGCT 382
DB	CCAACCTGACAGGACGCTGGCGCGGCTCGCGCGGAGGAGGCGGACGACATCGTAGGTGC 433
OY	ACTGCATTAACAGCAGTACAGCCCGCTGTGCTGCTGCTGATCGGGGCCCACTGTCAG 442
DB	TGAGGCGCCCGCAGCGGTACAGAGCCCGCGCGGTGCGCATCGGGCCCAACAGCACTC 493
OY	AGCTGCGCATGATGACGCGGCAAGTTCTTCAGTCTTCTCATGCCCAGGTCAGACTACG 502
DB	AGCTGGCCTTACCAACGAGCCCGCATCNNNNNGTCTTCTCTGTCGGAGATCACTACG 553
OY	GTGCTAGCATGAGATGCTGTGAGCCGCCGAGACCTTCCCTCTCTTCGACACCGT 560
DB	AAGCTCTCTGTGAGATGCTGAGACAGAAAGCGGTTCTTACCCCTCGTTCTTGGGACAT 611
RESULT 7	495 bp mRNA linear EST 28-FEB-2002
LOCUS	BMT06305
DEFINITION	UT-E-DMO-agh-a-22-0-UI r1 UT-E-DMO Homo sapiens cDNA clone
ACCESSION	BMT06305

VERSION	BM706305.1	GI:19019563
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE	1 (bases 1 to 495)	
JOURNAL	Bonaldo,M.B., Lennon,G. and Soares,M.B.	
MEDLINE	Normalization and subtraction: two approaches to facilitate gene	
COMMENT	discovery	
	genome Res. 6 (9), 791-806 (1996)	
	9704447	
	Contact: Soares, MB	
	Program for Rat Gene Discovery and Mapping	
	University of Iowa	
	451 Eckstein Medical Research Building Iowa City, IA 52242, USA	
	Tel: 319 335 8250	
	Fax: 319 335 9565	
	Email: msoares@blue.weeg.uiowa.edu	
	Tissue Procurement: Dr. Gregg Hageman	
	CDNA library preparation: Dr. M. Bento Soares, University of Iowa	
	CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa	
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa	
	Clone Distribution: Researchers may obtain clones from Research	
	Genetics (www.resgen.com).	
	Seq primer: M13 Reverse.	
FEATURES	Location/Qualifiers	
source	1..495	

```

/1bb host="DH10B (Life Technologies) (T1 phage resistant)
/notes="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-DW0 is a cDNA library containing the following
tissue(s): lens. The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT73-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CGATTACGA. This library
was created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."
```

VERSION	KEYWORDS	EST.
REFERENCE	AI562167.1	GI:4513512
AUTHORS	house mouse.	
	Mus musculus.	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus	
	1 (phases 1 to 298)	
	Maria,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,	

TITLE	The WashU-NCI Mouse EST Project 1999
JOURNAL	Unpublished (1999)
COMMENT	Contact: Marra M/WashU-NCI Mouse EST Project 1999 Mailbox: Marra@nci.nih.gov Address: National Cancer Institute, 301 N. Zeeb Road, Room 303A, Bethesda, MD 20892-4500

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:663147

This clone was previously sequenced on the 5' end only, this new data is from the 3' end

Possible reversed clone: similarity on wrong strand

Possible reversed clone: polyT not found

High quality sequence stop: 286.

FEATURES

SOURCE

```
1. 298 /organism="Mus musculus"  
/strain="NIH Swiss"  
/db_xref="taxon:10090"  
/clone_image:1260595"  
clone_11b="Stratagene mouse heart (#537316)"
```

```

      •
      /tissue.type="heart"
      /dev.stage="13 day embryos"
      /lab.host="SOLR (kanamycin resistant)"
      /note="Organ: heart; Vector: pBluescript SK-; Site: 1;
      EcoRI: Site: 2; XhoI: Cloned unidirectionally. Primer
      Oligo: dt_93 pooled NIH/Swiss 13 day embryo hearts.
      Average insert size: 1.0 kb, um-2AP XR vector; -5'
      adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor
      sequence: 5' CTCGACGTTTTTTTTTTTTTTTTT 3'
BASE COUNT      59 a      87 g      51 t
ORIGIN

```

Query Match	4.1%;	Score 104.8;	DB 9;	Length 298;
Best Local Similarity	63.5%;	Pred. No.-5e-09;		
Matches 160;	Conservative 0;	Mismatches 92;	Indels 0;	Gaps 0

OY	487	CCCCAGGTCACCTACAGGGCTAGCATGGAGCTGGTGAGACGCCCGGAGACCTTCCCTCC	548
Db	298	CCACAGGTCACACTATGAGGCGCAGCAGCTGATCTCTCATGTGGGAAGCCGCACTTCCGCTCC	239
OY	547	TTCTTCGCGACCGTGCCCGACGCACTGTGTGCAGCTGACGGCCGCGCGCGAGACTGTCCAG	606
Db	238	TTCTTCGCGCACACATCCCCAGGATGAAGTACAGAGCTGGTAAGTCATATGACGCTCTCCAG	179
OY	607	GAGTTCCGCGCTGGAACTGGGGTGGCCGCTGGGCGACGCGAGCTGAAGGGCGGCGCAAGGGC	666
Db	178	AGCTTCGGGCTGGGCTGTGATCTCCCTGTTTGGCAGCTATGGTGTACTACGGGCACTTGGC	119
OY	667	CTGAGCATCTTCTCGGCCCTGGCCGCGGACGCGGCGCATCTTGCAATCGCGCAGGAGGCTTG	726
Db	118	GTACAGGCGCTGGAGAGAGCTGGCCACTCCACAGGGGCACTGTGCTGCGCTTCAAGGAGCTG	59
OY	727	GTGCGCGCTGCCCC	738
Db	58	GTGCGCTCTCTCC	47

RESULT 9						
EG182090						
LOCUS		709 bp	mRNA	linear	EST 21-APR-2001	
DEFINITION	RG91954 Athersys RAGE Library Homo sapiens cdna, mRNA sequence.					
ACCESSION	EG182090					
VERSION	EG182090.1	GI:13703777				
KEYWORDS	EST.					
SOURCE	human.					

REFERENCE
AUTHORS
1 (bases 1 to 709) .
Harrington, J. J., Sherf, B., Rundlett, S., Jackson, P. D., Perry, R.,
Cain, S., Leventhal, G., Thornton, M., Ramachandran, P., Whittington, J.

TITLE	Creation of genome-wide protein expression libraries using random
JOURNAL	activation of gene expression
MEDLINE	Nat. Biotechnol. 19 (5), 440-445 (2001)
COMMENT	21227151
	Contact: Scott J. Cain
	Abbott
	Tec

3201 Carnegie Ave., Cleveland, OH 44115, USA
Tel.: 216 431 9900
Fax: 216 361 9596
Email: scaine@athersys.com
High quality sequence stop: 485.
ATGAGTCTGGTT

FEATURES

```
1. .709
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Athersys RAGE Library"
/cell_line="HT1080"
note="see 'Creation of Genome-wide Protein Expression
```

Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

Query Match	4.1%;	Score 104.6;	DB 12;	Length 709;
Best Local Similarity	52.2%;	Pred. No. 6e-09;		
Matches 258;	Conservative	0;	Mismatches 230;	Indels 6;
				Gaps 1

QY	244	GAGATCAACACCAAGTGGAGATCTCTGCGGGGCTGGCGGGGCTACAGACTTTTGAT	303
Db	54	GAGATGATCAACCAATTTCACACTCTTATCTGGAATCAAACTGGGGTATGAATCTATGAC	1133
QY	304	ACGTGCTCGGAGCTGTGTGTGGCCATGAAGCCAGCCTCATGTTCTGTGGCCAAAGCAGGC	363
Db	114	ACTGTACAGAAAGTCACAGATGGCAATGGACGCCACTCTGAGGTTCTTCTTAAATTCAAC	1733
QY	364	AGCGCGGACA-----TCGGCGGCTACTGGCAATCAACGGAGTACCAAGCCCGGGTGC	417
Db	174	TGCTCCAGAGAACTGTGGATTTAAAGTGTACTATTCCAGCTACATCAATCCAAAGGTTAAG	233
QY	418	GCTGTCTACGGGCCCACTGCTGACAGCTGGCCATGTGTCAACGGCAAGTTCTTACGCTTC	477
Db	234	GCTGTCAAGAGTTCTGTGGTATCTAGAAATAACTATGGCTGTCTCCAGAGATGTTGAATTTA	293
QY	478	TTTCTCATGCCCCCAAGGTCTACGTACGTGTCTACATGAGCTGCTGAGCGCCCGGGAGACC	537
Db	294	CAGTCTATGCCACAGGTGGGTTAATGAATCAACTGCACAAATAATCTCGATGCTCCAAATTCGC	353
QY	538	TTTCCCTTCCTTCTCCACACGTGTCCACAGACAGCGTGTGACGTGACGAGCGCGCGCGAG	597
Db	354	TTTCTCTCATTTTTCACGAGACTGTGCCCAAGTACCTTCATCAATAATTAAGCAATGTGCTCAC	413
QY	598	CTGTGACAGAGATTTCGCGTGAAGCTGGGTGGCGGCTGTGGACGACGACGAGATGCGC	657

Db	414	CGATTTCAGAAATTCGGTTTGGAACATCGGATTTGGCATCATACACAGATCATGATATGGA	473
Qy	658	CGGCAAGGGCCCTAGCATCTTCTCGGGCCCTGGCCCGGCGACGCGGCACTGCATCGGCGAC	717
Db	474	CAATTGGGCTCTTAACACTTTTATATTCAGAGCTGAAGCANATACGTGGCTACACTTC	533
Qy	718	GAGGGCGCTGTGTC 731	
Db	534	AAAGAGCTTCTTC 547	
RESULT 10			
LOCUS	AA907022	168 bp	EST 19-MAY-1998
DEFINITION	o192a08.s2 Soares_NFL_T_GRC_S1 Homo sapiens cDNA clone		
ACCESSION	AA907022		
VERSION	AA907022.1	GI:3042482	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 168)		
JOURNAL	NCI-GAP htcp://www.ncbi.nlm.nih.gov/ncigap.		
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)		
FEATURES	Source		
source	Location/Qualifiers		
	1..168		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:1505750"		
	/clone_11b="Soares_NFL_T_GRC_S1"		
	/lab_host="DH10B"		
	/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NDHL19W, testis NHT, and B-cell NCI-GAP-GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."		
ORIGIN	25 a	49 g	42 t
Query Match	4.0%	Score 102.2;	DB 9; Length 168;
Best Local Similarity	97.2%;	Pred. No. 1.4e-08;	
Matches 104; Conservative	0; Mismatches	3; Indels	0; Gaps 0;
Qy	1171	TTCTCTGCTACGACAGCTGTATATAGCTGGCCGAGCCCTCACAACACTTTCAGTGC	1230
Db	11	TTTTTTGCTACGACAGCTGTATATAGCTGGCCGAGCCCTCACAACACTTTCAGTGC	70
Qy	1231	AACGCTCAGGCTGCCCGCGCAGGACCCCGTGAAGCCCTGGCAGCT	1277
Db	71	AACGCTCAGGCTGCCCGCGCAGGACCCCGTGAAGCCCTGGCAGGT	117
RESULT 11			
LOCUS	CNS0506R/c	1038 bp	DNA linear GSS 26-JUL-2000

	DEFINITION
	Tetraodon nigroviridis genome survey sequence T7 end of clone 00M17 of library B from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION	AL315180.1 GI:9548068
VERSION	AL315180
KEYWORDS	GSS; genome survey sequence.
SOURCE	Tetraodon nigroviridis.
ORGANISM	Tetraodon nigroviridis. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetradontiformes; Tetradontidae; Tetraodon. 1 (bases 1 to 1038)
REFERENCE	Roeft Croallius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Minoke,P., Brotlier,F., Quetier,F., Saurin,W. and Weissensbach,J. Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence Nat. genet. 25 (2), 235-238 (2000) Nt. genet. 25 (2), 235-238 (2000)
AUTHORS	Nat. genet. 25 (2), 235-238 (2000)
TITLE	1 (bases 1 to 1038)
JOURNAL MEDLINE	2 (bases 1 to 1038)
PUBMED	Croalilus,H.R., Jallion,O., Dasilva,C., Ozouf-Costas,C., Fizames,C., Fischer.C., Bouneau.L., Billault.A., Quetier.F., Saurin.W., Bernot.A. and Weissensbach.J. Characterization and repeat analysis of the compact genome of the freshwater puiferfish Tetraodon nigroviridis Genome Res. 10 (7), 939-949 (2000)
REFERENCE	20359837
AUTHORS	10899143
JOURNAL	3 (bases 1 to 1038)
MEDLINE	Genoscope.
PUBMED	Direct Submission Submitted (12-APR-2000)
AUTHORS	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon. Location/Qualifiers
FEATURES	location/Qualifiers 1..1038 /organism="Tetraodon nigroviridis" /db_xref="taxon:99883" /clone="006M17" /clone_1id="B" /note="Genosope sequence ID : COAB006AG09CI -end : T7"
BASE COUNT	242 a 249 c 332 g 201 t 14 others
ORIGIN	
	Query Match 4.0%; Score 101.8; DB 17; Length 1038; Best Local Similarity 51.2%; Pred.No. Ze-08; Matches 371; Conservative 0; Mis-matches 337; Indels 17; Gaps 5;
Oy	1729 AGCCGCGGCCTTGTCGTGGCCTTTGGGCGTTGCCATCGGAGCAGC 1788
Dd	939 ACCCTGGGCACTCTCTGTGCTATTGACATCGCTCTTTCTTAAGAAGGGAACACT 880
Oy	1789 CCACCTGTTAAGCCTCGGGGGGGCCCCCTGGCCTCTTTGGCCTGTGTGCGGCGCTG 1848
Dd	879 CCGGGTGAAGAGCAGCGGGGACCCTTAGCC-AAGCANCTCTTCCTCCGTGATTC 821
Oy	1849 GTCTGCTCAGCTCTCTCTGTTCCCTGGCACCAAGCCAGTCCCTGCGCCAG 1908
Dd	820 AGCTCATCATCAGCGCGCCCTGTTTTGGGGCGAGCCCAACAACCTCCAAGTCAGAAAGCCCGC 761
Oy	1909 CAGCCCCTGTCACATCCCGCTCAGGGGGGAGCCCGACACACCTTCCTCGAGGCGGCC 1968
Dd	760 CAGGTGCTCTTCGGCATCAGCTTACTCTGTGTGTCTCTTCGATCTTGCTCAAGACTTGG 701
Oy	1969 GAGATCTTCGTGAAGTCAAAGTGGCTTGAAGTGGGAGAACGGAGTGGGCTG 2028
Dd	700 CAGATCTCGTAGAGCTTCCAGTTCAAGCCCGGGGCTGCAGAGAGCTGTAGAGAAGATCAT 641
Oy	2029 CGGGGGCCCTGGGCTGCTGTGTGTCTCTGCTGCAATGCTGTGTAGAGCTGCACACTGTGC 2088

/note="Organ: pooled; Vector: pT733-Ple (Pharmacia) with a modified polylinker; Stee1: Not I; Stee2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and six circles were made *in vitro*. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and clones: Soares NDH8 pool 1:

```
clone = ATG02.255200
/clone_1lb="Soares_NSF_F8_9W_OT_PA_P_S1"
```

/lab_host="DH10B"
 /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
 a modified polylinker; Site_1: Not I; Site_2: Eco RI;
 Equal amounts of plasmid DNA from five normalized
 libraries were mixed, and ss circles were made in vitro.
 Following HAP purification, this DNA was used as tracer in
 a subtractive hybridization reaction. The driver was
 PCR-amplified cDNAs from pools of 5,000 clones made from
 the same 5 libraries. The pools consisted of the following
 libraries and cloneIDs: Soares NbHSF pool 1:
 309384-110919, 333208-325895 Soares Nb2Hr pool 1:
 145032-147935, 147720-148103, 148872-149255, 15002 -
 150407, 151176-152327 Soares Nb2H8H-9w pool 1:
 758280-760583, 772104-774407 Soares NbHP pool 1:
 304776-306311, 320136-332823, 326580-326663 Soares NbHOT
 pool 1: 723720-726407, 739080-740999 subtraction by Benco
 Soares and M. Fatima Bonaldo."

[illegible]

OY 431 CCCACGCTCGAGAGCTCGCCCATGTCACCGGGCAATTCTTCAGCTTCTTCCCATGCCCC 490
 |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
 Db 317 CACACGGCTCAGGGCTCTCCACGGCAGTGGCAAACTCTGCTGGGGCTTTTTCATATCCCC 258
 OY 491 AGGTCAGCTACGGTGGCTAGCATGGAGCTGCTGAGGGCCCGGGAGACCTTCCCTCTTCT 550
 |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
 Db 257 AGGTCAAGTTATGCGCTCTCCAGCAGACACTCTTCAGCAACAAGATCAATCAAGTCTTCC 198
 OY 551 TCCGCACCGGTGCCACGCAACCGTGTGTCAGTCAACGGCCGCCCGGGAGGCTGTGCAAGAGT 610
 |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
 Db 197 TCCGAACCATCCCCCATATGATGATGACACACGGCCACTGCCATGGCAACAATCATCTGAGTATT 138
 OY 611 TCGCGTGGAACTCGGGTGGCCCGCTCGGGCAGCAGACAGTACGGCCGGCAGGGCCCTGA 670
 |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
 Db 137 TCCCGTGGAACTCGGGTGGGGCAAAATTGACGTGATGACGACATATGGGGCCGGGGGATTTG 78
 OY 671 GCATTTCTCGGCCCTGGCCCGCGGCGACGCGGCATCTGCATCGCGCAACGAGGGCCTGGTGC 730
 |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
 Db 77 AGAAATTCGAGAGGAGAGCTGAGGAAAGGAGATATCTGCATCGACTTCATGATGAATCATCT 18
 OY 731 CGCTGCCCCCGGCCGA 746
 |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||
 Db 17 CCCAGTACTCTGATGA 2
 |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | ||||

RESULT 14	
AM771483/c	
LOCUS	635 bp mRNA
DEFINITION	h57c05.x1 NC1 CGAP Kid1 Homo sapiens CDNA clone IMAGE:3031976 3'
	similar to SW:GSR_BOVIN_P53584 EXTRACELLULAR CALCIUM-SENSING RECEPTOR PRECURSOR ;, mRNA sequence.
ACCESSION	AM771483
VERSION	AM771483.1 GI:7703539
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 635)	Nci-CCAP	http://www.ncbi.nlm.nih.gov/ncicgap .		
	National Cancer Institute, Cancer Genome Anatomy Project (CCAP).			
	Tumor Gene Index			
	Unpublished (1997)			
	Contact: Robert Strausberg, Ph.D.			
	Email: cgapbs-r@mail.nih.gov			

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -40UP from GIBCO
High quality sequence: stop: 460.
Location/Qualifiers
1. .635

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3031976"
/clone_lib="NCI CGAP_Kid11"
/lab_host="DH10B"
/notes="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
hybridization, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDs 1323376-133911, 1436007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT      188 g      126 g      193 t
ORIGIN

```

[illegible]

RESULT 15
BG191110/c

1	279.4	10.9	2532	4	US-09-361-631-4	Sequence 4	Appl
2	284	11.1	2993	4	US-09-361-631-3	Sequence 3	Appl
3	218.6	8.5	2010	4	US-09-361-631-8	Sequence 8	Appl
4	173	6.8	3809	1	US-08-485-588-3	Sequence 3	Appl
5	173	6.8	3809	1	US-08-485-588-3	Sequence 3	Appl
6	173	6.8	3809	2	US-08-485-588-3	Sequence 3	Appl
7	173	6.8	3809	2	US-08-485-588-3	Sequence 3	Appl
8	173	6.8	3809	3	US-08-353-784-3	Sequence 3	Appl
9	173	6.8	3809	3	US-08-484-719B-3	Sequence 3	Appl
10	173	6.8	3809	4	US-08-546-998-2	Sequence 2	Appl
11	173	6.8	3809	4	US-08-484-159-3	Sequence 2	Appl
12	173	6.8	4000	2	US-08-687-289A-2	Sequence 2	Appl
13	173	6.8	5006	1	US-08-485-588-2	Sequence 2	Appl
14	173	6.8	5006	1	US-08-484-585-2	Sequence 2	Appl
15	173	6.8	5006	2	US-08-480-751-2	Sequence 2	Appl
16	173	6.8	5006	2	US-08-943-986-2	Sequence 2	Appl
17	173	6.8	5006	3	US-08-353-784-2	Sequence 2	Appl
18	173	6.8	5006	3	US-08-484-719B-2	Sequence 2	Appl
19	173	6.8	5006	4	US-08-546-998-1	Sequence 1	Appl
20	173	6.8	5006	4	US-08-484-159-2	Sequence 2	Appl
21	161.8	6.3	5275	1	US-08-485-588-1	Sequence 1	Appl
22	161.8	6.3	5275	1	US-08-484-751-1	Sequence 1	Appl
23	161.8	6.3	5275	2	US-08-943-986-1	Sequence 1	Appl
24	161.8	6.3	5275	2	US-08-353-784-1	Sequence 1	Appl
25	161.8	6.3	5275	3	US-08-484-719B-1	Sequence 1	Appl
26	161.8	6.3	5275	4	US-08-484-159-1	Sequence 1	Appl
27	161.8	6.3	5275	4	US-08-484-159-1	Sequence 1	Appl

28	152.8	6.0	2177	4	US-09-134-513-1	Sequence 1, Appl1
29	146.2	5.7	2188	5	PCR-US93-01642-1	Sequence 1, Appl1
30	144.8	5.7	4134	4	US-09-162-021B-1	Sequence 1, Appl1
31	144.8	5.7	4134	4	US-09-687-477-17	Sequence 17, Appl1
32	133.6	5.2	4131	1	US-08-485-586-4	Sequence 4, Appl1
33	133.6	5.2	4131	1	US-08-484-565-4	Sequence 4, Appl1
34	133.6	5.2	4131	2	US-08-480-751-4	Sequence 4, Appl1
35	133.6	5.2	4131	2	US-08-943-986-4	Sequence 4, Appl1
36	133.6	5.2	4131	3	US-08-353-784-4	Sequence 4, Appl1
37	133.6	5.2	4131	3	US-08-484-719B-4	Sequence 4, Appl1
38	133.6	5.2	4131	3	US-08-484-159-4	Sequence 4, Appl1
39	101.2	4.0	3384	2	US-08-687-289A-1	Sequence 1, Appl1
40	95.8	3.7	4095	5	PCR-US91-094322-18	Sequence 18, Appl1
41	93.8	3.7	4236	4	US-08-617-785-1	Sequence 1, Appl1
42	93.8	3.7	3451	4	US-09-641-318-1	Sequence 1, Appl1
43	93.8	3.7	3451	4	US-09-641-318-3	Sequence 3, Appl1
44	87.4	3.4	2436	5	PCR-US91-094322-20	Sequence 20, Appl1
45	77.8	3.0	2619	3	US-08-77-797A-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1

Sequence 4 Application US/09361631
 Patent No. 6383778
 GENERAL INFORMATION:
 APPLICANT: Zuker, Charles S.
 APPLICANT: Adler, Jon Elliot
 APPLICANT: Lindemeier, Juergen
 APPLICANT: The Regents of the University of California
 TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
 TITLE OF INVENTION: Involved in Sensory Transduction
 FILE REFERENCE: 02307E-088720US
 CURRENT APPLICATION NUMBER: US/09/361,631
 CURRENT FILING DATE: 1999-07-27
 EARLIER APPLICATION NUMBER: US 60/095,464
 EARLIER FILING DATE: 1998-07-28
 EARLIER APPLICATION NUMBER: US 60/112,747
 EARLIER FILING DATE: 1998-12-17
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 4
 LENGTH: 2532
 TYPE: DNA
 ORGANISM: Mus sp.
 FEATURE:
 OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) B4
 OTHER INFORMATION: nucleotide
 US-09-361-631-4

Query Match	11.1%	Score 284	DB 4	Length 2532
Best Local Similarly	48.6%	Pred. No. 6e-44		
Matches 1180, Conservative	0	Mismatches 1185	Indels 65	Gaps 12

OY	92	TGAAGGGGACTACGTGCTGGGGGGGGCGTTCCCGCGGGGAGAGCGGAGAG- - - - -G	14.5
Db	95	TGGCTGGGGACTACCTCTGGGGTGGCCCTTTACCTTCATGCCAACGTAGAGTGTCT	15.4
OY	146	CTGGGCTCCGACGGCGGACAGCGGCCACGACCCCTGTGTGCACGAGTTCTCCCAACG	20.5
Db	155	CTCACTTACGTACTCTGAGAGGTGGCCAAAGTGCATGAGTACATGTAAGGTGTGGCT	21.4
OY	206	GCCTGCTCTGGGCACGTGGCCATGAAATATGGCCGTGGAGAGAGATCAACACATGCGGATC	26.5
Db	215	ACAACCTCATG- - -CAGGCCCATGGATATCGCGCTGGGAGAAATCAACAACGTAGCTCTT	27.1
OY	266	TGGAGCCCCGGGCTCGGCCCTGGGGTACGACCCCTTGTGATGAGTGGTGGAGGCTGTGGTG	32.5
Db	272	TGCTGCCCGGGCTGTGCTCGGCTACAGAGATGTTGAGTGTGCT- - -ACCTTCACACA	32.8
OY	326	CCATGAAGCCCAAGCTCATGTTCTCTGGGCAAGGACGAGCGCGGACATCGCCGCTACT	38.5

RESULT 2

US-09-361-631-3

; Sequence 3, Application US/09361631

; Patent No. 6383778

; GENERAL INFORMATION:

APPLICANT: Zuker, Charles S.

APPLICANT: Adler, Jon Elliot

APPLICANT: Lindemeyer, Juerge

APPLICANT: The Regents of the University of California

1. TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor

TITLE OF INVENTION: Involved in Sensory Transduction

FILE REFERENCE: 02307E-0887200S

CURRENT APPLICATION NUMBER: US/

CURRENT FILING DATE: 1999-07-27

EARLIER APPLICATION NUMBER: US

EARLIER FILING DATE: 1998-07-28

EARLIER APPLICATION NUMBER: US

EARLIER AFFILIATION NUMBER: 1998-1

NUMBER OF SEC

NUMBER OF SE
SOFTWARE: PatSOFTWARE:
SEO ID NO 3; SEQ ID NO 3
LENGTH: 2002

LENGTH: 2993

```

; TYPE: DNA
; ORIGIN: Esthu

```

ORGANISM: R

Query Match	10.98;	Score 279.4;	DB 4;	Length 2993;
-------------	--------	--------------	-------	--------------

Best Local Similarity 48.38; Pred. No. 4.3e-43;

Matches 1189; Conservative 0; Mismatches 1211; Indels 63; Gaps 12;

QY	92	TGAGAGGGGAGTACGTCTGGGGGGGCTGTTCCTCCCTGGGGCAGAGCCGAGAG-----G	145
Db	107	TGGCCGGGGACTACTCTCTGGTGGTGGCTCTTAACTCCATGGCAACGTAAAGAGATCT	166
QY	146	CTGGCCTCCGACAGCCGGACACAGGCCACAGACCCCTGTGTACACAGAGTTCTCTCAACG	205
Db	167	CCCACTTCAGTACTCTGACAGGTGGCCAAAGTGCATATGATTCACCATGAAAGTCTTGGGCT	226
QY	206	GCTGCTCTGGGCACTGGCCATGAAAAATGGCCGTGGAGAGATCACACAGTCGATC	265
Db	227	ACAACCTCATG---CAAGCCATGCTGTTCCTGTGGAGGAGATCAACACTGTAGCTCCC	283
QY	266	TGCTGCCCGGGCTCGGCTGGGCTACGACCCCTTTGATGTACGTCTGGAGCCCTGTGGTGG	325
Db	284	TGCTACCCGGCGTCTCTGTCTGGCTACGAGATGGTGGATGTCTGT---ACCTTCCACA	340
QY	326	CCATGAAAGCCCAAGCCATGTTCTCTGGCCAAAGCAGGACCGCAGACATGCCCGCTACT	385
Db	341	ATATCCACCCCTGGGCTCTACTTCTCTGGCAAGAGAGCAGACATCCC---TCCCATCCTCA	397
QY	386	GCACCTACGCCATTAACGAGCCCGTGTGTGTGTGCTGTATCGGGCCCACTCGTCAGAGC	445
Db	398	AAGACTACAGCAATATATGCCCCACGCTGGGTGTGTCATTGTGGCCCCGACMACTGTAGT	457
QY	446	TGCGCATGTCACGGGAAGTTCTTCAGCTTCTCCATGACCCCAAGTCAGTCAGAGTGG	505
Db	458	CCGGCATTAACGTTGCCAACATTTCTCTCTATTTCTCATATCCACAGATACATACAGCG	517
QY	506	CTAGCATGAGAGCTGTGAGCGCCCGGAGAGACCTTCCCTCCTCTTCGACCGCTGCCA	565
Db	518	CCATCTCCGACAACTGAGGAGCAAGCGGACCTCTCCCTACATGATGCTACGCACAGTGCCA	577
QY	566	GCGACCGTGTCACTGACAGGCGCCCGGAGCGTGTGTCAGAGATTTGGGTGGAACTGGG	625
Db	578	GCGGCACCCACACATGAGAGCGCAATGGTGCAGCTGATGTTCACTTCMAATGGAACCTGGA	637
QY	626	TGGCGGCTCTGGGACGAGACACATGACGGCCGACAGGCGCTGAGCATTT---CTCGG	682

D	638	TTGATGCTCTGTTAGCGAGACAGATTACGGCCCGGAGAACGCAACTGTTTGAGCCAGC	697
Q	663	CCCTGGCCGGCGACACGGGCATCTGCATGCGCGACAGAGGGCCCTGGTGCCCGCTGG	742
D	698	GCTTAGCCAAACAGAGGACATTCGATTGCGCTTCAGAGAGGTTCTGCCCTACCTAGAT	757
Q	743	CCGATGA-----CTGCGGCGCTGGGGAAGGTGAGGACGTCCTCGACCAAGGGA	790
D	758	CCAGCCAGGTATAGAGTCCGAGAGAGACACTGGACAACATCTCGGACAAGCTGC	817
Q	791	ACCAAGAGCGGTCAGAGTGGTGTCTGTGTTGCGCTCCCTGCACGCCCGCCACGCGCTT	850
D	818	GGCGAGCTCGGCGCGGTGTGGTGGTGTCTTCGCGCGAGCTGAGCCGTATAGCTTCT	877
Q	851	TCAACTACAGCATAGACAGCAGGCTCTCGCCCAAGGTGGTGGTGGCCACGACGCTTGGC	910
D	878	TTCAAGAGAGTCTCCGGTGGAACTTCAACGGGTTTGTGGATGCGCTGTGAGTCTGGG	937
Q	911	TGACCTTGACCTGGTATGGATGGGCGTGGCCGGCATGGCCAGATGGGACAGGCTTGGCT	970
D	938	CTATCGAACCCAGTTCTGTCAATTAACCTCAACGAGTGGCGCCACACGGGTACTTTCTGGGCG	997
Q	971	TCCTCCAGAGGGGTGCCACGCTGACGAGTGTCCCCAGTACGTGAAGACGACCTGGGCC	1030
D	998	TCACCATTCACAGAGGTTCCATCCCTGGCTTCACTGATCGATCCGAGT-----GCGC	1047
Q	1031	TGGCCACCGACCCGCGCTTCTGCTCTGCTCCCTGGGCGAGAGGAGACAGGTTCTGGAGAGG	1090
D	1048	CGTGACAAAGCCAGGGTATCCCGTGGCTTAACAGCAGCAACTTCGGACAGCTGGCAACCG	1107
Q	1091	ACGTGGTGGGCGCACGCTGCCCGCAGTGTGACTGCATCAACGCTGGAGAACGTAGCCGAG	1150
D	1108	GACTGTGACGCGCTCTTGAACACA-----CCAAAGCTTTCAACAACATCTTATAC	1159
Q	1151	GGCTAAATCACACACACAAACGTTCTGTCTGTCTACGCACTGTGATAGCTGGCCAGGCC	1210
D	1160	TTTGGGGGAGCGGTGTCTTACAGCGGTGACTCGGCACTTTACCGGGGGCCATYGCC	1219
Q	1211	TGCACAAACCTCTCACTGACAGCCCTCAAGGCTGACCCCGAGAGAACCCCGTAAGCCCT	1270
D	1220	TCACAGACGCTCTGGGTGTAAACCGGTCGGCGTACACAGCAAA-----GGTATACCGCT	1276
Q	1271	GGCAGCTCTCGGAGAACATGTACAACTGACCTTCAAGTGGGCGGCGCTCCGCTGGCGT	1330
D	1277	GGCAGCTACTGAGGAATCTGGACGCTCAACTTCAACGCTCTCGGTAAACCGGCTCTTCT	1336
Q	1331	TCGACAGCAGCGGAAGGTGGACATGGAGTACACCTGAACGCTGAGGTTGGGTTGGCGAGGCT	1390
D	1337	TTGACCAACAGGGGACATGCGGATGCTCTTGGACATCATCAAGTGGCAGTGGACCTGA	1396
Q	1391	CAGTGCCCGAGCTCCACAGCGTGGGCGAGTTCAACGGCAGCCTCAG-----ACAGAGC	1444
D	1397	GCCAAATCCCTTCCAAAGCATGCGCTCTCTATCTCCACACAGCAAGAGGCTAACTTACA	1456
Q	1445	GGCTGAATATCCGTTGGCACAACGTCTGACACAACAGAGCCCGTGTCCGGTCTTGGCGGC	1504
D	1457	TTTAACAATGTCTCTGTGTATACCCCCAAACAACGGTCTCCTCTTCCATGTGTGTCCAAGA	1516
Q	1505	AGTGGCAGAGAGGCGCAGGTGGCGCGGCTCAAGGGGTTCCATCTCCGTCTCTACGATCTG	1564
D	1517	GCTGGCACCCAGGGCAATGAAAAAGTCTGTGGGCTCCACCCCTTGTTCCTTCAGATGCT	1576
Q	1565	TGGACTGCGAGGCGGACGCTACCGGCAAAACCCAGACAGCATCGGCTGACACTTGTGTG	1624
D	1577	TGGATTGTATGTCAGAGCACTACTCAACCGCTCACAGATAGTTTAATCTGTCTCT	1636
Q	1625	GCCAGAGAT---GATGGGCCCCGGAGGAGACAGACAGCGTCTTCCGCGCAGGTGTGGCT	1681
D	1637	GCCCGGGTTCATGTGGTCTTACAGAGACATCATCTGTCTTCCACGCGCGGCGCTTACT	1696
Q	1682	TCCTGGCATGGGCGGAGCGGCTGTGCTGTCTGTCTCTGTGACTGACCTGGCGCTGG	1741
D	1697	TCCTTGATGAGAGAGATGCCCAATGTGGTGGCCATATGAGGTGGCTGGCGCTCTCT	1756

QY 1742 GCGTTGCTGGGCTGCTTTGGGGCTGTTGTTACCATCGGAGACCCACTGGTTCAGG 1801
 Db 1757 TCAGTACACTGGCCATTCCTTTTCAATCTTGAGACATTTCCAGACACCACCTGGTGGCT 1816
 QY 1802 CCTGGGGGGGGCCCTGGCTGCTTGGCTTGGCTGCTGGCTGGCTGGCTGGCTGGCT 1861
 Db 1817 CGGGCGGG 1876
 QY 1862 TCCTCTGCTGCTGCTGGCCAGCCCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1921
 Db 1877 TGCCCGGTATATGAGGG 1936
 QY 1922 ACCTCCGCTGACGGGGCTGCTGACACACTTCTGCTGACGGGGGGGGGGGGGGGGGG 1981
 Db 1937 CCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1996
 QY 1982 AGTGAAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2041
 Db 1997 TCTTCAGATGGCCAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2056
 QY 2042 CCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2101
 Db 2057 CCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2116
 QY 2102 TGCCCTTCCCGCCGGAGGTGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2161
 Db 2117 TGCGCCACCATCATCAACCCCATTTGGCCGAGCCGACCCGGATGACCCCAATCATGATCC 2176
 QY 2162 ACTGCGGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2218
 Db 2177 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2236
 QY 2219 TGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2278
 Db 2237 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2296
 QY 2279 GTGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2338
 Db 2297 AAGCCAACTCATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2356
 QY 2339 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2398
 Db 2357 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2416
 QY 2399 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2458
 Db 2417 TCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2476
 QY 2459 AGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2518
 Db 2477 ACCCGGAGCGGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2536
 QY 2519 AGA 2521
 Db 2537 AGA 2539

RESULT 3
 US-09-361-631-8
 : Sequence 8, Application US/09361631
 : Patent No. 6383778
 : GENERAL INFORMATION:
 : APPLICANT: Zuker, Charles S.
 : APPLICANT: Adler, Jon Elliot
 : APPLICANT: Lindemeier, Juergen
 : APPLICANT: The Regents of the University of California
 : TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
 : FILE REFERENCE: 02307E-088720US
 : CURRENT APPLICATION NUMBER: US/09/361,631
 : CURRENT FILING DATE: 1999-07-27
 : EARLIER APPLICATION NUMBER: US 60/095,464

: EARLIER FILING DATE: 1998-07-28
 : EARLIER APPLICATION NUMBER: US 60/112,747
 : EARLIER FILING DATE: 1998-12-17
 : NUMBER OF SEQ ID NOS: 10
 : SOFTWARE: Patent In Ver. 2.0
 : SEQ ID NO 8
 : LENGTH: 2010
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : OTHER INFORMATION: human G-protein coupled receptor (GPCR) B4
 : OTHER INFORMATION: nucleotide sequence
 : US-09-361-631-8
 Query Match 8.5%; Score 218.6; DB 4; Length 2010;
 Best Local Similarity 46.8%; Pred. No. 5,9e-32;
 Matches 941; Conservative 0; Mismatches 1014; Indels 54; Gaps 6;
 QY 494 TCAGCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 553
 Db 2 TCACCTACAGGCGCCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 61
 QY 554 GCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
 Db 62 GTACACACACCGAGCGCCGAGCACCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
 QY 614 GCTGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 673
 Db 122 GCTGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 181
 QY 674 TCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 733
 Db 182 TGCTTGGCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
 QY 734 TGCCCGGCGAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 781
 Db 242 CACTGCAAGCCACCAACACATGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
 QY 782 ACCAGGTGAACAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 841
 Db 302 ACAAGGTGAACAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361
 QY 842 AGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 901
 Db 362 ACCACTTCTTCAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421
 QY 902 AGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 961
 Db 422 AGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481
 QY 962 TGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1021
 Db 482 TCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541
 QY 1022 ACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1081
 Db 542 GC-----CCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 591
 QY 1082 TGGAG 1141
 Db 592 TGCAACAG 651
 QY 1142 TGAGCGAG 1201
 Db 652 CTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 703
 QY 1202 CCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1261
 Db 704 CCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 760
 QY 1262 TGAAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1321
 Db 761 TCTACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 820

TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 373..3606
 OTHER INFORMATION:
 US-08-485-588-3

Query Match 6.8%; Score 173; DB 1; Length 3809;
 Best Local Similarity 51.1%; Pred. No. 1.5e-23;
 Matches 497; Conservative 0; Mismatches 455; Indels 21; Gaps 3;

QY 79 CAGCACTTAGGATGAGGAGGACTAGCTGCTGGGGGGCTGTTCCCTGGCGAGGCC 138
 DB 442 CAGGAGGCCCAAGAGGGGGACATATCTTGGGGGGCTTCTCCATTCATTTTGA 501
 QY 139 GAGGAGGCT---GGCCTCCGAGCCGAGCCAGCCAGCCAGCCCTGTGTACACAGTTC 195
 DB 502 GTACAGCTTAAGATTCAGATCTCAATCAAGCCGGAGTCTGTGGAATGTATCAGGTAT 561
 QY 196 TCCCTAAGCGGCTGCTGTGGGCACTGGCCATGAAATGCGCTGAGAGATCAACAC 255
 DB 562 AATTCGGTGGGTTTCGCTGTACAGGCTATGATTTGGCATAGAGAGATTAACAGC 621
 QY 256 AAGTCGATCTGCTCCCGGGCTCGCTGGGCTACGACTCTTGTATGCTGCTGGAG 315
 DB 622 AGCCACGCCCTTCTCCCACTTACGCTGGGATGAGATATTGACACTTGCAACAC 681
 QY 316 CCTGTGGGCGCCATGAGCCGCTCATGTTCTGCGC-----CAGGCGAGGACGCCG 369
 DB 682 GTTCTAAGGCTTGGAGCCACCTGAGTTTGTGCTCAAAACAAATGATTTCTTG 741
 QY 370 GACATCGCCGCTTACCACTACAGCAGATACAGCCCGCTGCTGCTGCTATCGGG 429
 DB 742 AACCTGATGAGTCTGCAACTGCTCAGAGCAGATTCCTTACGATTTGCTGTTGGA 801
 QY 430 CCCCACCTGTCAGAGCTGCGCATGTCACCGGCAAGTTCTTCACTTCTTCTATGCC 489
 DB 802 GCAATGCTGTCAGGCTGCTCCAGCGAGTGCAATCTGCTGGGGCTTCTTACATCCC 861
 QY 490 CAGGTCACTACGCTGCTAGCATGAGCTGCTGAGCGCGCGGAGACCTTCCCTCTC 549
 DB 862 CAGGTCACTTATGCTCTCTCAGACAGACTCTCAGCAACAAATTCATTTAAGTCTTC 921
 QY 550 TTCCGCACTGCTGCGCAGACCTGCTGAGCTGACGGCGCGCGGAGCTGCTCAGAG 609
 DB 922 CTCGAGCACTTCCCAATGATGAGCAGCAGGCTGCTCAAGCAGATCATGAGAT 981
 QY 610 TTGGGCTGGAAGCTGGGGCGCCCTGGGCGACAGCAGATGAGCTGCGCGCAGGCTG 669
 DB 982 TTCCGCTGGAAGCTGGGGCGCAATTCAGCTGATGACGCTATGAGGCGCGCGGAT 1041
 QY 670 AGCATCTTCTGCGCCCTGCGCGCGCAGCGGCACTGATCGCGCAGAGGCGCTG 729
 DB 1042 GAGAAATTCGAGAGAGAGAGAGAGATTCGATCGACTTCAGTGAATCATC 1101
 QY 730 CCGCTGCGCGCGGCTGCTGCTGCGGCTGGGAGGTGAGGAGCTGCTGACACAGTG 789
 DB 1102 TCCAGTACTCTGATGA-----GGAAGAGATTCAGCATGATGATGAGGTGAT 1149
 QY 790 AACGAGAGAGCGCTGAGGTGCTGCTGCTGCTGCTGCTGCGCGCGCGCGCGCTC 849
 DB 1150 CAAATTCAGAGCGCAAGATCTGCTGCTTCTCCAGTGGCCCAAGATCTTGAACCCCTC 1209
 QY 850 TTCATACAGCATCAGCAGAGCTCTGCGCCCAAGGTGAGGTGGGCGAGGAGCGCTG 909
 DB 1210 ATCAAGAGATTTGCTCCGCGCATATACGCGCAGATCTGCTGCGAGGAGGCTGG 1269
 QY 910 CTGACCTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 969
 DB 1270 GCGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1329
 QY 970 TTCTCCAGAGGGGCTGCCAGCTGACGAGTTCCTCCCAAGTACGTAAGACGACCTG 1029

DB 1330 TTCTCTGAGAGCTGGGCGAGATCCAGGCTTCCGGGAATCTGAGAAAGTCCATCC 1389
 QY 1030 CTGGCCACCGCACC 1042
 DB 1390 AGGAGTCTGTCC 1402

RESULT 5 US-08-484-565-3

Sequence 3, Application US/08484565
 Patent No. 5763569

GENERAL INFORMATION:

APPLICANT: Edward M. Brown
 APPLICANT: Steven C. Hebert
 APPLICANT: James E. Garrett, Jr.
 TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 MOLECULES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Lyon & Lyon
 STREET: First Interstate World Center
 STREET: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: USA
 ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,565
 FILING DATE: 7 June, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA: including application
 PRIOR APPLICATION DATA: described below: 9

APPLICATION NUMBER: 08/353,784

FILING DATE: 9 December, 1994
 APPLICATION NUMBER: PCT/US/94/12117

FILING DATE: 21 October, 1994
 APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994
 APPLICATION NUMBER: U.S. 08/141,248

FILING DATE: 22 October, 1993
 APPLICATION NUMBER: U.S. 08/009,389

FILING DATE: 23 February, 1993
 APPLICATION NUMBER: U.S. 08/017,127

FILING DATE: 12 February, 1993
 APPLICATION NUMBER: U.S. 07/934,161

FILING DATE: 21 August, 1992
 APPLICATION NUMBER: U.S. 07/834,044

FILING DATE: 11 February, 1992
 APPLICATION NUMBER: U.S. 07/749,451

FILING DATE: 23 August, 1991
 ATTORNEY/AGENT INFORMATION:

NAME: Hebert, Sheldon O.
 REGISTRATION NUMBER: 38,179

REFERENCE/DOCKET NUMBER: 213/006
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440

TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
 LENGTH: 3809 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA

FEATURE:
NAME/KEY: CDS
LOCATION: 373..3606
OTHER INFORMATION:
US-08-480-751-3

Query Match 6.88; Score 173; DB 2; Length 3809;
Best Local Similarity 51.18; Pred. No. 1.5e-23;
Matches 497; Conservative 0; Mismatches 455; Indels 21; Gaps 3;

79 CAGCACTTGGATGAGGAGGAGTACGTGTGGGGGGGCGTGTCCCGGGGGAGGCC 138
DB 442 CAGGAGACCCCAAGAGAGAGGAGGAGATATCTTGGGGGCTCTTCTATTCATTTTGA 501
QY 139 GAGGAGGCT--GGCTCCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 195
DB 502 GTACGACTTAAGATCAAGATCTCAATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAT 561
QY 196 TCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 255
DB 562 AATTCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 621
QY 256 AAGTCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 315
DB 622 AGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 681
QY 316 CCGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 369
DB 682 GTTCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 741
QY 370 GAGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 429
DB 742 AACCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 801
QY 430 CCCCAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 489
DB 802 GCACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 861
QY 490 CAGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 549
DB 862 CAGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 921
QY 550 TTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 609
DB 922 CTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 981
QY 610 TTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 669
DB 982 TTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1041
QY 670 AGCATCTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 729
DB 1042 GAGAAATTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1101
QY 730 CCGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 789
DB 1102 TTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1149
QY 790 AACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 849
DB 1150 CAAATTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1209
QY 850 TTTCACTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 909
DB 1210 ATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1269
QY 910 CTGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 969
DB 1270 GCCGAGCTTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1329
QY 970 TTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1029
DB 1330 TTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1389

QY 1030 CTGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1042
DB 1390 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1402

RESULT 7
US-08-943-986-3
Sequence 3, Application US/08943986
Patent No. 5962314

GENERAL INFORMATION:

APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/943,986
FILING DATE: 03-Oct-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/484,565
FILING DATE: 7-June-1995
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991

ATTORNEY/AGENT INFORMATION:

NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO. 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 3809 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

FEATURE: NAME/KEY: CDS

US-08-353-784-3

Query Match 6.88; Score 173; DB 3; Length 3809;
 Best Local Similarity 51.18; Pred. No. 1.5e-23;
 Matches 497; Conservative 0; Mismatches 455; Indels 21; Gaps 3;

QY 79 CAGCACTTGAAGAGAGGAGTACTGCTGGGGGGGCTTCCCTGGGGAGAGCC 138
 DB 442 CAGCGAGCCCAAGAGAGGAGGACATTTCTTGGGGGGGCTTCTTATTCATTTTGA 501
 QY 139 GAGGAGGCT---GGCCTCCGAGCGGAGAGGAGCCAGAGCCCTGTGTGACACCGTTC 195
 DB 502 GTAGAGCTAAGATCAAGATCTCAATCAAGGCGGAGCTGTGGAATGATTCAGGAT 561
 QY 196 TCTCAAGAGGCTGCTGTGGGCACTGGCCATGAATAAGCCGTGGAGAGATCAACAC 255
 DB 562 AATTTCGTGGGTTTCGTGGTTACGGCTATGATTTGGCATAGAGAGATTAACAGC 621
 QY 256 AAGTCGATCTGCTCCCGGGCTGGCTGGGCTAGCACTTGTGATACGTCTGAG 315
 DB 622 AGCCAGGCTTCTCCCACTTGCAGCTGGGATACAGGATATTTGACACTTGCACAC 681
 QY 316 CCGTGGTGGCCATGAAGCCAGCCCTCATGTTCCGCGC-----CAAGGAGGAGCCGC 369
 DB 682 GTTCTAAGGCTTGAAGCCACCTGATTTGTCTCAAAACAAATGATTTCTTG 741
 QY 370 GACATGCGGCTTACTGCACTACAGCAATACAGCCCGCTGCTGCTCATTCGG 429
 DB 742 AACCTTGATGATTTCTGCACTGCTCAGAGCAATTCCTCTAGCATTTGTTGGGA 801
 QY 430 CCCCCTGTGAGAGCTGGCCATGCTCACCGCAAGTTCTTACGTTTCTCATGCC 489
 DB 802 GCACTGGCTCAGGGGCTCCACAGGCACTGCAATCTGCGGGCTCTTACATTTCC 861
 QY 490 CAGCTCAGCTACGCTGATGAGATGAGCTGCTGAGCGCCCGGAGAGACCTTCCCTTC 549
 DB 862 CAGCTCATTTATGCTCTCTCCAGCAGACTCTCTCAGCAACAGAAATTCATTTTC 921
 QY 550 TTCGCAACCGTGGCCAGGAGCGTGTGAGTGAAGCGCGCGCGAGACTGCTGAGAG 609
 DB 922 CTCGCAACCATCCCAATGATGAGACAGGCACTGCCATGCGCAGCATATTCAGAT 981
 QY 610 TTCGCTGGAAGTGGGTGGCGGCTGGGAGAGCAGTACGCGCGGAGGCGCTG 669
 DB 982 TTCGCTGGAAGTGGGTGGGACATTCAGCTGATGACGATGAGGCGCGGCGGAT 1041
 QY 670 AGCATCTCTGCGCCCGCGGCGGAGCTGATCGAGCGGCGGAGAGGCGCTGCTG 729
 DB 1042 GAGAAATTCAGAGAGAGAGCTGAGAAAGGATATCTGATCGACTTCACTGATC 1101
 QY 730 CCGCTGCCCCGTGCGGATGCTCGCGGCTGGGAGAGTGAAGAGCTCTGACACAGG 789
 DB 1102 TCCGAGTACTGTGATGA-----GGAGAGATTCAGCATGTGTGAGAGTGT 1149
 QY 790 AACCAAGACAGCGTGAAGTGTGCTGTCTGCTCCCTGCGTGCAGCGCCGCGCCCTC 849
 DB 1150 CAAATTCAGGCGCAAGGATATCTGTTTCTCAGTGGCCGAGATCTTGAAGCCCTC 1209
 QY 850 TTCAACTACAGATCAGAGAGGCTCTCGCCCAAGGTGGGTGGGAGAGGCGCTG 909
 DB 1210 ATCAAGAGATTTGCGCGGCAATATCACGGGCAAGATCTGGCTGCGAGGAGCTG 1269
 QY 910 CTGACCTCTGACTGTGATGAGGCTGCGCGGATGCGCCAGATGAGGAGCGTGTG 969
 DB 1270 GCGAGCTCTCTCCGATGAGCGATGCTGATCTTCAAGGATGAGGAGCGATTTGA 1329
 QY 970 TTCTCTCAGAGGAGGCTGAGAGATTTCCCAAGTGAAGAGAGAGAGAGAGCTG 1029
 DB 1330 TTCTCTCAGAGGAGGCTGAGAGATTTCCCAAGTGAAGAGAGAGAGAGAGCTG 1389
 QY 1030 CTGGCAGCGAGC 1042
 DB 1390 AGGAAGTCTGTCC 1402

RESULT 9

US-08-484-719B-3

Sequence 3, Application US/08484719B

Patent No. 6031003

GENERAL INFORMATION:

APPLICANT: Edward F. Nemeth, Edward M.

APPLICANT: Brown, Steven C. Hebert,

APPLICANT: Bradford C. Van Wagenen,

APPLICANT: Manuel F. Balandrin

APPLICANT: Forrest H. Fuller, Eric G.

APPLICANT: Delmar, Scott T. Moe

TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: First Interstate World Center

STREET: Suite 4700

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: MS Word

SOFTWARE: FastSeq for Windows Version 3.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,719B

FILING DATE: 7 June, 1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/353,784

FILING DATE: 9 December, 1994

APPLICATION NUMBER: PCT/US/94/12117

FILING DATE: 21 October, 1994

APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994

APPLICATION NUMBER: U.S. 08/141,248

FILING DATE: 22 October, 1993

APPLICATION NUMBER: U.S. 08/009,389

FILING DATE: 23 February, 1993

APPLICATION NUMBER: U.S. 08/017,127

FILING DATE: 12 February, 1993

APPLICATION NUMBER: U.S. 07/934,161

FILING DATE: 21 August, 1992

APPLICATION NUMBER: U.S. 07/834,044

FILING DATE: 11 February, 1992

APPLICATION NUMBER: U.S. 07/749,451

FILING DATE: 23 August, 1991

ATTORNEY/AGENT INFORMATION:

NAME: Douglas C. Murdock

REGISTRATION NUMBER: 37,549

REFERENCE/DOCKET NUMBER: 213/007

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 3809 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

FEATURE:

NAME/KEY: CDS

LOCATION: 373..3606

US-08-484-719B-3

QY 256 AAGTGGATCTGCTCCCGGGCTGGGCTAGCACTCTTTGATACGTCTGGAG 315
 Db 622 AGCCAGAGCCCTTCTCCCACTTGGATACAGATATTTGACACTTGCACAC 681
 QY 316 CCTGTGGGCGCATGAAAGCCCTCATGTTCTGGC-----CAAGCAGGCGCC 369
 Db 682 GTTCTTAAGGCGCTTGAAGCCAGCTGATTTGTTGCTCAAAACAAATTTG 741
 QY 370 GACATCGCCGCTTGTGCACTATACAGCAATACAGCCCGGCTGCTGATCGGG 429
 Db 742 AACCTGATGATGCTGCACTGCTCAGACACATTCCTCTACATGCTGTGGGA 801
 QY 430 CCCCAGCTGAGAGCTGCGCATGCTACCGCAAGTCTTCACTTCTCTATGCC 489
 Db 802 GCACTGAGCTCAGAGGCTCTCCACGCGCATGGCAATCTGCGGGCTTCTTACATTC 861
 QY 490 CAGGTCACTACGCTGCTGATGAGTGTGAGCCCGGAGACCTTCCCTCTTC 549
 Db 862 CAGGTCACTATGCTCTCTCCAGCAGACTCTCTCAAGCAAGATCAATTCAGTCTTC 921
 QY 550 TTCCGACCGCTGCGCCAGGCGCTGCTGACGCTGACGCGCGCGGAGCTGCTGAGAG 609
 Db 922 CTCGGAACCATCCCAATGATGACAGCCACTGCTGATGCGCATCATGAGTAT 981
 QY 610 TTCCGCTGGAAGTGGGCGCCCTGGGCGAGCAGCAGATACGCGCGGAGGCTG 669
 Db 982 TTCCGCTGGAAGTGGGCGCCCTGGGCGAGCAGCAGATACGCGCGGAGGCTG 1041
 QY 670 AGCATCTCTGCGGCTGCGCGCGGCGAGCAGCAGATACGCGCGGAGGCTG 729
 Db 1042 GAGAAATTCGAGAGAGAGTGAAGAAAGATATCTGATGACTTGAATGACTATC 1101
 QY 730 CCGCTGCGCCGCTGCGGATGCTGCGGCGGAGGAGTGCAGTCTCTCACAGCTG 789
 Db 1102 TCCAGTACTCTGATGA-----GGAAGATCCAGCATGTGTGAGAGTAT 1149
 QY 790 AACCAAGACAGCTGACAGTGTGCTGTTCCGCTCCGTCAGCGCCGACCCCTC 849
 Db 1150 CAAATTCAGAGCGCAAGATATGCTGTTCTCCAGTGGCCCGAGATCTTGAAGCCCTC 1209
 QY 850 TTCACTACATGATGAGCAGAGGCTGCGCCCAAGGATGGGCGGAGGAGGCGCTG 909
 Db 1210 ATCAAGAGATTTTCGCGCGCAATATCACGCGCAAGATCTGCGGAGGAGGCTG 1269
 QY 910 CTGACCTCTGACCTGATGAGGCTGCGCGCATGCGCCAGATGCGCGAGCGTGTGGC 969
 Db 1270 GCCAGCTCTCCCTGATGCGCATGCTGCTGATCTTCCAGCGTGGTGGGAGCATTTGA 1329
 QY 970 TTCTCCAGAGAGGCGCCAGCTGACAGATTTCCCGCAGTACGTAAGAGCCAGCTG 1029
 Db 1330 TTCCGCTGGAAGGCTGGGCGAGATCCAGGCTTCCGGAATTCCTGAAAGAGTCCATCC 1389
 QY 1030 CTGGCCAGCGAGC 1042
 Db 1390 AGGAAGTCTGTC 1402

RESULT 11
 US-08-484-159-3
 : Sequence 3, Application US/08484159
 : Patent No. 6313146
 : GENERAL INFORMATION:
 : APPLICANT: Bradford C. Van Wagenen
 : APPLICANT: Manuel F. Balandrin
 : APPLICANT: Eric G. Del Mar
 : APPLICANT: Edward F. Nemeth
 : TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 : TITLE OF INVENTION: MOLECULES
 : NUMBER OF SEQUENCES: 20
 : CORRESPONDENCE ADDRESSES:
 : ADDRESSEE: Lyon & Lyon
 : STREET: First Interstate World Center
 : STREET: Suite 4700

STREET: 633 West Fifth Street
 CITY: Los Angeles
 STATE: California
 COUNTRY: USA
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: FASTSEQ
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,159
 FILING DATE: 7 June, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 PRIOR APPLICATION DATA: including application
 PRIOR APPLICATION DATA: described below: 9
 APPLICATION NUMBER: 08/353,784
 FILING DATE: 9 December, 1994
 APPLICATION NUMBER: PCT/US/94/12117
 FILING DATE: 21 October, 1994
 APPLICATION NUMBER: U.S. 08/292,827
 FILING DATE: 23 August, 1994
 APPLICATION NUMBER: U.S. 08/141,248
 FILING DATE: 22 October, 1993
 APPLICATION NUMBER: U.S. 08/009,389
 FILING DATE: 23 February, 1993
 APPLICATION NUMBER: U.S. 08/017,127
 FILING DATE: 12 February, 1993
 APPLICATION NUMBER: U.S. 07/934,161
 FILING DATE: 21 August, 1992
 APPLICATION NUMBER: U.S. 07/834,044
 FILING DATE: 11 February, 1992
 APPLICATION NUMBER: U.S. 07/749,451
 FILING DATE: 23 August, 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Hebert, Sheldon O.
 REGISTRATION NUMBER: 38,179
 REFERENCE/DOCKET NUMBER: 214/101
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3809 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 373..3606
 OTHER INFORMATION:
 US-08-484-159-3

Query Match 6.8%; Score 173; DB 4; Length 3809;
 Best Local Similarity 51.1%; Pred. No. 1.5e-23;
 Matches 497; Conservative 0; Mismatches 455; Indels 21; Gaps 3;
 QY 79 CAGCACTTAGATGAAGGGGAGTACGTGCTGGGGGCTGTTCCTCCGCGGAGGCC 138
 Db 442 CAGCGAGCCCAAAAGAGGGGACATATCTTGGGGGCTCTTCTCATTCATTTGGA 501
 QY 139 GAGGAGGCT---GGCGTCCGAGCGGAGCAGCGGCCAGAGCCCTGTGTGACAGGTC 195
 Db 502 GTACAGCTAAAGATCAAGATCTAAATCAAGCGCGGAGTGTGTGATATATAGGAT 561
 QY 196 TCCTCAAGGCGCTGCTGGGCACTGCGCATGAAATGCGCGTGGAGAGATCAACAC 255
 Db 562 AATTCGCTGGCTTCCCTGTACAGGCTATATTTGCAAGAGAGATTAACAGC 621
 QY 256 AAGTGGATCTGCTCCCGGGCTGGGCTAGACACTCTTTGATACGTCTGGAG 315

```

Db 622 AGCCAGCCCTTCTCCAACTGACGTGGATACAGATATTGACACTTGCACAC 681
Qy 316 CCTGTGTGCCATGAGAGCCAGCTCTATGTTCTGCGC-----CAAGGACAGGACCCG 369
Db 682 GTTCTTAAGGCGCTTGAGAGCACCCTGAGTTGTGCTCAAAACAAATTAATGATCTT 741
Qy 370 GACATCGCGCGCTACTGAACTACAGCAGATACAGCCCGCTGTGCTGTATCGGG 429
Db 742 AACCTTGATGATCTTGAACTGCTCAGAGACATTTCCCTAGCATTTGTTGTGGGA 801
Qy 430 CCCCAGCTGAGAGAGCTGCGCATGATGACCGGAGTTCTTCACTTCTCTGATGCCC 489
Db 802 GCACTGAGCTGAGGCGCTGCTCCAGCGCATGAGCAATTCGTCGGGCTTTTACATTTCC 861
Qy 480 CAGCTGACATAGGCTGTAGCATGAGAGCTGTGAGCGCCCGGAGACCTTCCCTCTTC 549
Db 862 CAGGTCACTTATGCTCTCTCCAGCAGACTCCTCAGCAACAAGAAATCAATTAAGTCTTC 921
Qy 550 TTCCGACACGTCGCCAGGACCGGTGTGAGCTGACGGCGCGCGGAGCTGTGAGAG 609
Db 922 CTCGGAACCATCCCAATGATGAGCAGCAGCCACTGCGCATGCGACATATGAGAT 981
Qy 610 TTCGCTGGAAGTGGGTGCGCCCTGCGGACGAGCAGATAGCGCGGAGGCGCTG 669
Db 982 TTCGCTGGAAGTGGGTGCGGACATGATGAGCATGATGAGCGGCGCGGAGAT 1041
Qy 670 AGCATCTTCTGCGCGCTGCGCGGACGCGGACATGCTGATGCGGACGAGGCGCTGTG 729
Db 1042 GAGAAATTTCCAGAGAGAGCTGAGAGAGGATATCTGATGATGATGATGATGAT 1101
Qy 730 CCGTGCCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
Db 1102 TCCAGTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1149
Qy 790 AACGAGACAGCTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 849
Db 1150 CAAATTCACAGCGCCCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1209
Qy 850 TTCAACTACAGATGAGAGAGCTGCGCCCAAGGTGTGCTGCTGCTGCTGCTGCTG 909
Db 1210 ATCAAGGAGATGCTGCGCGCAATATACAGGAGATGCTGCTGCTGCTGCTGCTG 1269
Qy 910 CTGACCTGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 969
Db 1270 GCCAGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1329
Qy 970 TTCTCCAGAGGCGTCCGACGATGATGATGATGATGATGATGATGATGATGATG 1029
Db 1330 TTCGCTGAGAGGCTGGGACATGATGATGATGATGATGATGATGATGATGATG 1389
Qy 1030 CTGGCAGCAGC 1042
Db 1390 AGGAGTCTGTCC 1402

```

RESULT 12
US-08-687-289A-2
Sequence 2, Application US/08687289A

Patent No. 5981195
GENERAL INFORMATION:
APPLICANT: Fuller, Forrest H.
APPLICANT: Krapcho, Karen J.
APPLICANT: Hammerland, Lance G.
TITLE OF INVENTION: CHIMERIC RECEPTORS AND METHODS FOR
TITLE OF INVENTION: IDENTIFYING COMPOUNDS ACTIVE AT
TITLE OF INVENTION: METABOTROPIC GLUTAMATE RECEPTORS AND
TITLE OF INVENTION: THE USE OF SUCH COMPOUNDS IN THE
TITLE OF INVENTION: TREATMENT OF NEUROLOGICAL DISORDERS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon

```

STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Fastseq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687, 289A
FILING DATE: July 25, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,526
FILING DATE: July 26, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Walburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 220/004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4000 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-687-289A-2

```

Query Match 6.8%; Score 173; DB 2; Length 4000;
Best Local Similarity 51.1%; Pred. No. 1.5e-23;
Matches 497; Conservative 0; Mismatches 455; Indels 21; Gaps 3;

```

Qy 79 CAGCACTTAGAGTAGAGGAGGACTGCTGCTGCGGGGCTGTGCTCCCTGGCGAGGCC 138
Db 110 CAGCAGGCCCAAGAGAGGAGGACATTTCTTGGGGGCTCTTCTATTCATTTTGA 169
Qy 139 GAGGAGGCT---GGCTCGGACGCGGACAGCGGCCACAGCCCTGTGTGACCGAGTTC 195
Db 170 GTAGAGCTAAAGATCAAGATCTCAATCAAGCCGAGCTGTGTGAATGTATCGAT 229
Qy 196 TCTCAAAAGGCTGCTGCTGCGGACATGCGCATGAATGAGCCGTGTGAGATCAACAAC 255
Db 230 AATTTCGCTGGGTTTCGCTGTGATGAGCTATGATATTTGCCATGTAGAGATTAACAGC 289
Qy 256 AAGTGTGATGTGCTGCGCGGCTGCGCTGAGCTAGCAGCTTTTGTATGATGCTGGAG 315
Db 290 AGCCAGGCTTCTTCCCACTGACGCTGGGATACAGATATTTGACACTTGCAACACC 349
Qy 316 COTGTGTGGCATGAACCCAGCCTCATGTTCTGCGC-----CAAGGACAGGACGCCG 369
Db 350 GTTCTAAGGCTTGGAAGCACCCTGAGTTTGTGCTCAAAATAAATTAATTTTGG 409
Qy 370 GACATCGCGCGCTACTGCACTACAGCAGATACAGCCCGCTGTGCTGCTCATCGGG 429
Db 410 AACCTTGATGATTTCTGCACTGCTCAGAGCAGCATTCCTGTAGATGTGTGTGGGA 469
Qy 430 CCCCAGCTGTAGAGCTGCCATGATGATGATGATGATGATGATGATGATGATGATG 489
Db 470 GCACTGAGCTAGAGGCTTCCAGCGCAGTGTGCAATTTGCTGGGCTTCTTACATTTCC 529
Qy 490 CAGGTCACTAGGCTGTAGCATGAGAGCTGTGAGAGCGCGGAGACCTTCCCTCTTC 549
Db 530 CAGGTCACTATGCTCTCTCAGAGCAGATCTCTCAGCAACAAGATTAATTAAGTCTTC 589
Qy 550 TTCCGACAGCTGCGGACGAGCAGCTGTGTGAGCTGAGCGCGCGGAGAGCTGTGAGAG 609

```

Db 590 CTCGCAACATCCCAATGATGACACAGCCACTGCTCATGCGAGATCATCATGATAT 649
 QY 610 TTGGGTGAACATGGGTGGGCGCCCTGGGCGAGACAGATGAGGCGGCGAGGCGCTG 669
 Db 650 TTCCGCTGGAATGAGGTGGGCGACATGTCAGCTGATGCGATATGGGCGCGGAGATT 709
 QY 670 AGCATCTCTGCGGCGCGCGCGAGCAGCGGATCTGATGCGGCGAGAGGCGCTGATG 729
 Db 710 GAGAAATTCGAGAGGAGAGGAGGAAAGGATATCTGATCATGCACTTCAGTGAATCATC 769
 QY 730 CCGTGGCGGCGCGGAGTACTCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 789
 Db 770 TCCGAGATCTGATGA-----GGAAGAGATCCGATGCTGATGAGAGTAT 817
 QY 790 AACGAGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 849
 Db 818 CAAATTCAGAGCGCAAGATCATGATGATTTCTTCAGTGGCCAGATCTTGAGCCCTTC 877
 QY 850 TTCACTACAGCATCAGCAGCAGGAGGCTCTCGCCCAAGGTGAGGAGGAGGAGGAGGAG 909
 Db 878 ATCAAGAGATGTCGCGGCGAATATCAGGCGCAAGATGCTGCTGGCAGAGGAGGAGG 937
 QY 910 CTGACCTCTGACCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 969
 Db 938 GCGAGCTCTCTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 997
 QY 970 TTCTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1029
 Db 998 TTGCTCTGAGAGGCTGGGAGATCCAGAGGCTTCCGAGGATTCGAGAGGAGGAGGAG 1057
 QY 1030 CTGGCCACCGAGC 1042
 Db 1058 AGGAGATCTGTC 1070

RESULT 13
 US-08-485-588-2
 Sequence 2, Application US/08485588
 Patent No. 5688938
 GENERAL INFORMATION:
 APPLICANT: Edward M. Brown
 APPLICANT: Steven C. Hebert
 APPLICANT: Forrest H. Fuller
 APPLICANT: James E. Garrett, Jr.
 TITLE OF INVENTION: CALCULUM RECEPTOR-ACTIVE
 TITLE OF INVENTION: MOLECULES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: First Interstate World Center
 STREET: Suite 4700
 STREET: 633 West Fifth Street
 CITY: Los Angeles
 STATE: California
 COUNTRY: USA
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: FASTSEQ
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/485,588
 FILING DATE: 7 June, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 PRIOR APPLICATION DATA: including application
 PRIOR APPLICATION DATA: described below: 9
 APPLICATION NUMBER: 08/353,784
 FILING DATE: 9 December, 1994
 APPLICATION NUMBER: PCT/US/94/12117
 FILING DATE: 21 October, 1994
 APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994
 APPLICATION NUMBER: U.S. 08/141,248
 FILING DATE: 22 October, 1993
 APPLICATION NUMBER: U.S. 08/009,389
 FILING DATE: 23 February, 1993
 APPLICATION NUMBER: U.S. 08/017,127
 FILING DATE: 12 February, 1993
 APPLICATION NUMBER: U.S. 07/934,161
 FILING DATE: 21 August, 1992
 APPLICATION NUMBER: U.S. 07/834,044
 FILING DATE: 11 February, 1992
 APPLICATION NUMBER: U.S. 07/749,451
 FILING DATE: 23 August, 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Hebert, Sheldon O.
 REGISTRATION NUMBER: 38,179
 REFERENCE/DOCKET NUMBER: 213/005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ. ID NO.: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5006 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 436..3699
 OTHER INFORMATION:
 US-08-485-588-2

Query Match 6.88; Score 173; DB 1; Length 5006;
 Best Local Similarity 51.18; Pred. No. 1,6e-23;
 Matches 497; Conservative 0; Mismatches 455; Indels 21; Gaps 3;
 QY 79 CAGCACTAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 138
 Db 505 CAGGAGACCCAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 564
 QY 139 GAGGAGGCT--GGCTCCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGGAG 195
 Db 565 GTAGCACTTAAGATCAAGATCAATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 624
 QY 196 TCCCAAGAGGCGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 255
 Db 625 AATTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 684
 QY 256 AAGTGGATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 315
 Db 685 AGCCAGCCCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 744
 QY 316 CCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 369
 Db 745 GTTCTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 804
 QY 370 GAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 429
 Db 805 AACCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 864
 QY 430 CCCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 489
 Db 865 GCAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 924
 QY 490 CAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 549
 Db 925 CAGTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 984
 QY 550 TTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 609
 Db 985 CTCGCAACATCCCAATGATGAGACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1044

```

OY 610 TTCGGCTGAACTGGGTGGCCGCTGGGACAGCAGAGTACGGCCGAGGCGCTG 669
Db 1045 TTCGCTGGAAGTGGGTGGGCAATTCAGCTGATGACGATAGGGGGGCGGAGATT 1104
OY 670 AGCATCTTCTGGCCCTGGCCGCGCAGCGCATCTGCATCGCCAGAGGGGCTGGTG 729
Db 1105 GAGAAATTCGAGAGAGAGCTAGAGAAAGGATATCTGCATGACATTGAGTAATCATC 1164
OY 730 CCGCTGCCCCGTCGAGTACTCGCGGCTGGGGAAGGTGCAGAGAGCTCTGCACCAAGTG 789
Db 1165 TCCCAGTACTGATGA-----GGAAAGATTCAGCAGATGTGTAGAGTGATT 1212
OY 790 AACGAGAGAGGCTGAGGTGTGCTGCTTCGCTCCGCTGACGCGCCGACCGCCCTC 849
Db 1213 CAAATTCACAGGCGCAAAATCATCTGCTTTCTTCAGTGGCCGCAATCTTGAGCCCTC 1272
OY 850 TTCAACTACAGCATCAGCAGCAGGCTCTCCGCAAGGTGTGGGTGGCCAGCGAGCTG 909
Db 1273 ATCAAGAGATGTCCGGGCAATATCAAGGCAAGATCTGGCTGGCCAGGAGGCTG 1332
OY 910 CTGACCTCTGACCTGTATGGGGCTGCCCGCATGCGCCAGATGGCAGGCTGCTTGGC 969
Db 1333 GCGAGCTCTCCCTGATGCCATGCTCAGTACTTCACAGTGTGGCGGCAACCATTTGGA 1392
OY 970 TTCTCCAGAGGGGTGCCAGCTGCAGAGTTCCCGCAGTACGTAGAGCGACCTGGCC 1029
Db 1393 TTGCTCTGAAGGCTGGGCGACATCCAGGCTTCGCGGAATTCCTAGAGAGTCCATCCC 1452
OY 1030 CTGGCCACCGACC 1042
Db 1453 AGAAGTCTGTCC 1465

```

RESULT 14
US-08-484-565-2
Sequence 2, Application US/08484565
Patent No. 5763569

GENERAL INFORMATION:

APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: James E. Garbett, Jr.
TITLE OF INVENTION: CALCULON RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,565
FILING DATE: 7 June, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9

APPLICATION NUMBER: 08/353,784

FILING DATE: 9 December, 1994

APPLICATION NUMBER: PCT/US/94/12117

FILING DATE: 21 October, 1994

APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994

APPLICATION NUMBER: U.S. 08/141,248

```

FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5006 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 436..3699
OTHER INFORMATION:
US-08-484-565-2

```

Query Match 6.8%; Score 173; DB 1; Length 5006;

Best Local Similarity 51.1%; Pred. No. 1,6e-23;

Matches 497; Conservative 0; Mismatches 455; Indels 21; Gaps 3;

```

OY 79 CAGCACTTAGATGAAGGGGAGTACGTGCTGGGGGGGCTGTCCCTGGCGAGGCC 138
Db 505 CAGCAGGCCCAAGAGAGGGGACATTTATCTTGGGGGGCTTTCTTATTCATTTTGA 564
OY 139 GAGGAGGCT---GGCTCCGACGCGGACGCGCCAGCCCTGTGTGCACGAGTTC 195
Db 565 GTAGCAGCTAAAGATCAGATCTCAATCAAGCCGAGTCTGTGAATGTATCAGGTAT 624
OY 196 TCCTCAAGCGCTCTCTGGGCACTGGCCATGAAAGGCGCGTGGAGAGATCAACAC 255
Db 625 AATTTCCGTGGTTTGGCTGTACAGGCTATGATATTTGCCATAGAGAGATTAACAGC 684
OY 256 AAGTCGATCTGCTCCGCGGCTGGCGCTGGGCTGACGACCTTTGATACGTGTCGGAG 315
Db 685 AGCCAGGCCCTTCTCCCACTGACGCTGGGATACAGATATTTGACACTTGCACACAC 744
OY 316 CCGTGTGGCCCATGAGCCCGACCTCATGTCTGGC-----CAGGCGAGCGCCGC 369
Db 745 GTTTTAAGGCTTGAAGCCACCTGAGTTTGTGCTCAAAACAAATATGATTTTGTG 804
OY 370 GACATCGCGGCTACTGCACTACACGACATACACGACGCGCGCTGCTGCTGATCGGG 429
Db 805 AACCTTGATGATTTCTGCACTGCTCAGACGACATTCCTCTACGATTTGCTGTGGTGA 864
OY 430 CCCCAGCTGAGAGCTGCGCATGCTACCGGCAAGTCTTCAGTTCCTCTCAATGCGCC 489
Db 865 GCAATGCTGCTGAGGCTCTCCACGCGCATGGAATTCGTCTGGGCTCTTCTACATTTCC 924
OY 490 CAGTCACTGAGTGTGACATGAGTGTCTGACGCGCCGCGGAGACCTTCCCTCTTC 549
Db 925 CAGTCACTTATGCTCTCTCCAGCAGACTCTCACAACAAGAAATCAATTAAGCTTTC 984
OY 550 TTCGCGACCGTGGCCAGGACCGTGTGACGTGACGCGCGCGCGGACCTGCTGAGAG 609
Db 985 CTCGCAACATCCCAATGATGAGCACCAGGCGCTGCAATGCGAGATCATCATGATAT 1044
OY 610 TTCGCTGAACTGGGTGGCGCCCTGGGACGACGACGATAGCGCGGAGGCGCTG 669

```

```

Db 1045 TTCGCTGGAAGTGGGGGAGCAATTGACACCTGATGACGACTATWGGCGCGCGGGGATT 1104
QY 670 AGCAATCTTCGCGCCCGCGGCAAGCGGCAATCTGACATGCGGCAAGAGGCGCTGGTG 729
Db 1105 GAGAAATTCGAGAGAGAGCTGAGAGAAAGGATATCTGATCGACTTGTAGTGAATCATTC 1164
QY 730 CCGCTGCCCGCTGCGGATGACTCGCGCTGGGGAAGTGCAGAGAGTCTCTCCACCAGGTG 789
Db 1165 TCCCAAGTACTGTATGA-----GGAAGAGATCCAGCATGTGTGTAGAGGTGATT 1212
QY 790 AACCAAGACAGCGTGCAGAGTGTCTGTCTGTCTGCTCCGTCGACAGCCCGCCAGCCCTC 849
Db 1213 CAAATTCACAGCGCAAGTATGTGTGTCTTCACAGTGGCCCAAGATCTTGACCCCTTC 1272
QY 850 TTCAACTACAGCATGAGAGAGGCTCGCGCCCAAGGAGTGGGAGGCGGAGGCGCGG 909
Db 1273 ATCAAGAGATATTCGCGCGCAATATACAGGAGATCTGCTGGCGGAGGAGGCGCTGG 1332
QY 910 CTGACCTCTGACCTGTGATGAGGCTGCGCGGCAATGCGGCAATGAGGAGCGTGTGGC 969
Db 1333 GCCAGCTCTCCCTGATCGGCTGATCGCTGATCTTCCAGCGTGGTGGGCGGACCATTTGA 1392
QY 970 TTCTCCAGAGGGGTGCCCACTGACAGAGTTCCTCCAGTACGTGAAGACGACCTGGCC 1029
Db 1393 TTGCTCTGAAAGCTGGGCGAGATCCAGGCTTCGCGGAATTCCTGAAGAAAGTGCATCCC 1452
QY 1030 CTGGCCACCGACG 1042
Db 1453 AGGAAGTCTGTCC 1465

```

RESULT 15
US-08-480-751-2
Sequence 2, Application US/08480751
Patent No. 585684

GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,751
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/35,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248

```

FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/004
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5006 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 436..3699
OTHER INFORMATION:
US-08-480-751-2

```

Query Match 6.8%; Score 173; DB 2; Length 5006;
Best Local Similarity 51.1%; Pred. No. 1.6e-23;
Matches 497; Conservative 0; Mismatches 455; Indels 21; Gaps 3;

```

QY 79 CAGCACTTAGATGAGAGGGGACTACGTGCTGGGGGCTGTTCCTCCGCGGAGGCC 138
Db 505 CAGCGAGCCCAAAAGAGGGGACATTAATCTTGGGGGCTCTTCCATTCATTTTGA 564
QY 139 GAGGAGCT--GGCCCTCCGAGCGGAGCAAGCGCCAGAGCCCTGTGTGACAGGTTGC 195
Db 565 GTACAGCTAAGATCAAGATCTAAATCAAGCGGAGTCTGTGAGATGATAGGATAT 624
QY 196 TCTCAACAGCCCTGCTGGGCACTGGCCATGAAATGGCCGTGAGAGATCAACAC 255
Db 625 AATTCGCTGGGCTTGCCTGCTGATGAGGATATTTGCAAGAGAGATTAACAGC 684
QY 256 AAGTGGATCTGCTGCGGGGCTGCGCTGGGCTAGACCTCTTTGATPAGCTGCGAG 315
Db 685 AGCCACGCCCTTCTTCCCACTTACGCTGGGATACAGATATTTGACACTTGCAACAC 744
QY 316 CCTGTGTGGCAATGAGAGCCAGCCCATGTCCTGCG-----CAGGAGAGAGCGGC 369
Db 745 GTTCTAAGGCTTGGAGCCACCTGATTTGTCTCAAAACAAATTTGATCTTTG 804
QY 370 GACATCGCCGCTTACCTCACTACAGCAGTACAGCCCGGTGTGCTGCTGATCGGG 429
Db 805 AACCTTATGATGCTCTGCACTGCTGAGAGCACTTCCCTCTAGATGCTGTGTGGGA 864
QY 430 CCCCACCTGTCAGAGCTGCCAGTGTACCGGCAAGTTCTTCACTTCTTCATGCCC 489
Db 865 GCAACTGGCTCAGGCTCTCCAGCGAGTGGCAAAATCTGCTGGGGCTCTCTCATATCC 924
QY 490 CAGGTGAGTACGCTGTAGATGAGTGTGAGCGCCCGGGAGACCTTCCCTCTTC 549
Db 925 CAGGTGAGTACGCTCTCTCAGAGACTCTCTCAAGAAATCAATTCATGCTTTTC 984
QY 550 TTGCGACCGGCGCCAGCAGCGTGTGACGTGACGCGCCCGCGGAGCTGTGACGAG 609
Db 985 CTCGAGACCATCCCAATGATGAGACACAGCCACTGCTGATGAGACATCATCGATAT 1044
QY 610 TTGCGCTGGAAGTGGTGGCGCGCTTGGGAGGAGAGAGATAGGCGGAGGCGCTG 669

```



```
Db 1045 TTCCGCTGGAACTGGGTGGGCGCAATTGCAAGCTGATGACAGACTATGCGGCGCGGGGGAATT 1104
QY 670 AGCATCTTCTGCGCCCTGGCCGCGGACGCGGCATCTGCATCGCGCACGAGGGCTGGTG 729
Db 1105 GAGAAATTCCGAGAGAGAAAGGATATCTCATCGACTTCAGTGAATCATTC 1164
QY 730 CCGCTGCCCGGCGGATGACTCGGGGCTGGGAGATGCAAGAGTCTCTGCACCAAGTG 789
Db 1165 TCCGATCTGTGATGA-----GGAGAGATCCAGCATGTGTAGAGGTGATT 1212
QY 790 AACGAGACGCGGTGACAGGTGCTGCTGCTTTCGCTCCGTGACAGCGCCGACGCGCTC 849
Db 1213 CAAAATTCCAGCGCAAGTCAATGCTGTTTCTCCAGTGGCCGACATCTTGAAGCCCTC 1272
QY 850 TTCAACTACAGCATACGACAGAGCTCTCGCCCAAGGTGTGGTGCCAGCGAGGCTTG 909
Db 1273 ATCAAGAGATGTCTCGGCGCAATATCAAGGCAAGATCTGGCTGGCCAGGAGGCTTG 1332
QY 910 CTGACCTCTGACCTGTGATGCGGCTGCGCCGATGCGCCAGATGCGACGCTGCTTGGC 969
Db 1333 GCCAGCTCTCTCCCTGATGCGCATGCTCTAGTACTTCCAGGTGTGGCGCACCATTTGA 1392
QY 970 TTCTTCCAGAGGGGTGCCAGCTGCACGAGTTTCCCCAGTACGTGAAGACGACCTGGCC 1029
Db 1393 TTGCTGTGAAGGCTGGGCGAGATCCAGAGCTTCCGGGAATTCCTGAAGAGGTCCATCCC 1452
QY 1030 CTGGCCACCGACC 1042
Db 1453 AGGAAGTCTGTCC 1465
```

Search completed: May 23, 2003, 18:24:46
Job time : 153.053 secs

THIS PAGE BLANK (USPTO)

QY 121 TTCCCTTGGGCGAGGCGGAGAGGCTGCTCCGACGCCGACAGGCGCCAGAGCCCT 180
Db 121 TTCCCTTGGGCGAGGCGGAGAGGCTGCTCCGACGCCGACAGGCGCCAGAGCCCT 180
QY 181 GTGTGACACAGGTTCTCTCAACGCGCTGCTGTGGGCACTGGCATGAAATGCGCTG 240
Db 181 GTGTGACACAGGTTCTCTCAACGCGCTGCTGTGGGCACTGGCATGAAATGCGCTG 240
QY 241 GAGGAGATCAACAAACAAAGTGTGCTGCGCGGCGGCTGGGCTAGAGCTCTTT 300
Db 241 GAGGAGATCAACAAACAAAGTGTGCTGCGCGGCGGCTGGGCTAGAGCTCTTT 300
QY 301 GATAGCTCTGGAGGCTGTGTGGGCACTGAAGCCGACCTCATGTTCTTGGCCAAAGCA 360
Db 301 GATAGCTCTGGAGGCTGTGTGGGCACTGAAGCCGACCTCATGTTCTTGGCCAAAGCA 360
QY 361 GGGAGCGCGGACATGCGCGCTACTGCACTACACGAGTACAGAGCCCGGTGTGCTG 420
Db 361 GGGAGCGCGGACATGCGCGCTACTGCACTACACGAGTACAGAGCCCGGTGTGCTG 420
QY 421 GTCATCGGGCCCACTGCTCAGAGCTGSCATGSCATGSCATGSCATGSCATGSCATG 480
Db 421 GTCATCGGGCCCACTGCTCAGAGCTGSCATGSCATGSCATGSCATGSCATGSCATG 480
QY 481 CTCAATGCCCCAGTCACTACGCTACGCTGCTAGATGAGTGTGAGCGCCGGAGACCTTC 540
Db 481 CTCAATGCCCCAGTCACTACGCTACGCTGCTAGATGAGTGTGAGCGCCGGAGACCTTC 540
QY 541 CCCCTCTTCTTCCGACCGCTGCGGACCGGACCGGCTGTGCACTGACGCGCGCGGAGCTG 600
Db 541 CCCCTCTTCTTCCGACCGCTGCGGACCGGACCGGCTGTGCACTGACGCGCGCGGAGCTG 600
QY 601 CTGCAAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 660
Db 601 CTGCAAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 660
QY 661 CAGGCGCTGAGCATCTTCTGCGCCCTGCGCGCGGACCGGCACTGTCATGCGCGACAG 720
Db 661 CAGGCGCTGAGCATCTTCTGCGCCCTGCGCGCGGACCGGCACTGTCATGCGCGACAG 720
QY 721 GGGCTGGGCGGCTGCGCGCGGAGTACGCGGCGGAGGAGGAGTGTGAGTGTGAGTGTG 780
Db 721 GGGCTGGGCGGCTGCGCGCGGAGTACGCGGCGGAGGAGGAGTGTGAGTGTGAGTGTG 780
QY 781 CACCAAGTGAACAGAGAGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 781 CACCAAGTGAACAGAGAGGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 841 CAGGCGCTTCAACTACATACATACATACATACATACATACATACATACATACATAC 900
Db 841 CAGGCGCTTCAACTACATACATACATACATACATACATACATACATACATACATAC 900
QY 901 GAGGCGTGGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 960
Db 901 GAGGCGTGGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 960
QY 961 GTGCTGGCTTCTCAACAGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1020
Db 961 GTGCTGGCTTCTCAACAGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1020
QY 1021 CACCTGGCTTCTCAACAGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1080
Db 1021 CACCTGGCTTCTCAACAGAGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1080
QY 1081 CTGAGAGGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1140
Db 1081 CTGAGAGGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1140
QY 1141 GTGAGCGCAGGCGTAAATCAACACAGAGCTTCTGTCTACAGAGCTGTATAGCGTG 1200
Db 1141 GTGAGCGCAGGCGTAAATCAACACAGAGCTTCTGTCTACAGAGCTGTATAGCGTG 1200

QY 1201 GCCAGGCGCTGACACACACTTTCAGTACAGAGGCTTACAGGCTGCGCGCAGAGACCC 1260
Db 1201 GCCAGGCGCTGACACACACTTTCAGTACAGAGGCTTACAGGCTGCGCGCAGAGACCC 1260
QY 1261 GTGAAGCCCTTGGACAGCTTCTGGAGAACATGTACAACTGACCTTCCACAGTGGGCGCTG 1320
Db 1261 GTGAAGCCCTTGGACAGCTTCTGGAGAACATGTACAACTGACCTTCCACAGTGGGCGCTG 1320
QY 1321 CCGCTGGGCTTGCAGACAGAGGAGAAAGTGGACATGAGTACAGAGCTGAGCTGAGCTG 1380
Db 1321 CCGCTGGGCTTGCAGACAGAGGAGAAAGTGGACATGAGTACAGAGCTGAGCTGAGCTG 1380
QY 1381 TGGCAGGAGCTCAAGTCCAGAGCTCCACAGCTGAGGCTGAGGCTCAACAGGAGCTGAGGACA 1440
Db 1381 TGGCAGGAGCTCAAGTCCAGAGCTCCACAGCTGAGGCTGAGGCTCAACAGGAGCTGAGGACA 1440
QY 1441 GAGGCGCTGAAGATCCGCTGGCAGACGCTGTGACAAACAGAAAGCCGCTGCTGCTG 1500
Db 1441 GAGGCGCTGAAGATCCGCTGGCAGACGCTGTGACAAACAGAAAGCCGCTGCTGCTGCTG 1500
QY 1501 CCGCAGTGCAGAGAGGCGCAGAGTGGCGCGGCTGCAAGGAGGCTTCCACTGCTGCTAGAC 1560
Db 1501 CCGCAGTGCAGAGAGGCGCAGAGTGGCGCGGCTGCAAGGAGGCTTCCACTGCTGCTAGAC 1560
QY 1561 TGTGTGACTGCGAGGCGGAGGAGCTACCGGCAAAACAGAGACATGCTGACCTTT 1620
Db 1561 TGTGTGACTGCGAGGCGGAGGAGCTACCGGCAAAACAGAGACATGCTGACCTTT 1620
QY 1621 TGTGTGAGATGAGTGTGCTCCCGGAGGAGAGACAGCTTCCGCGGAGAGTCTG 1680
Db 1621 TGTGTGAGATGAGTGTGCTCCCGGAGGAGAGACAGCTTCCGCGGAGAGTCTG 1680
QY 1681 TTCTGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
Db 1681 TTCTGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740
QY 1741 GGCCTGTGTGCTGCTTGTGGGCTGTGCTGACATGCGGAGAGCCACTGCTGCTG 1800
Db 1741 GGCCTGTGTGCTGCTTGTGGGCTGTGCTGACATGCGGAGAGCCACTGCTGCTG 1800
QY 1801 GCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
Db 1801 GCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
QY 1861 GTCTCTGCTTCCCTGAGCAGCCAGCCCTGCGGAGCTGAGGCGCAGAGCCCTTGTCC 1920
Db 1861 GTCTCTGCTTCCCTGAGCAGCCAGCCCTGCGGAGCTGAGGCGCAGAGCCCTTGTCC 1920
QY 1921 CACCTCCGCTACAGGAGGCTGCTGAGCAGCTTCTGAGGCGCGCGAGATCTTGTG 1980
Db 1921 CACCTCCGCTACAGGAGGCTGCTGAGCAGCTTCTGAGGCGCGCGAGATCTTGTG 1980
QY 1981 GAGTACAGAGCTGCTGAGAGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2040
Db 1981 GAGTACAGAGCTGCTGAGAGTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2040
QY 2041 GCGTGGT 2100
Db 2041 GCGTGGT 2100
QY 2101 GTGGCTTCCCGCGGAGGAGTGTGACGAGACATGTGCTGAGGAGGAGGAGGAGGAGGAGG 2160
Db 2101 GTGGCTTCCCGCGGAGGAGTGTGACGAGACATGTGCTGAGGAGGAGGAGGAGGAGGAGG 2160
QY 2161 CACTGAGGAGAGCTGCTGAGGAGTGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
Db 2161 CACTGAGGAGAGCTGCTGAGGAGTGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
QY 2221 GCGTCTGCTGCTTCTGAGGAGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
Db 2221 GCGTCTGCTGCTTCTGAGGAGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
QY 2281 GCCGCTGAGCTACCTTGGCAGCTGAGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340

Db 2281 GCGCGGGGCTACCTTGGCCATGCTGGGCTACTTCACTACCTGGGCTTGTGCCCC 2340
QY 2341 CTCCTGGCCAAATGTGAGTGTCTCTCAGAGCCCGCTGAGATGGGCGCCTCTGCTC 2400
Db 2341 CTCCTGGCCAAATGTGAGTGTCTCTCAGAGCCCGCTGAGATGGGCGCCTCTGCTC 2400
QY 2401 TGTGCTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
Db 2401 TGTGCTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
QY 2461 CCAGGGCTCAACACCCCGAGTCTCTCTGAGAGGGGCGCTGAGGATGCCCAAGCCAG 2520
Db 2461 CCAGGGCTCAACACCCCGAGTCTCTCTGAGAGGGGCGCTGAGGATGCCCAAGCCAG 2520
QY 2521 AATGACGGGAACACAGGAATCAGGGGGAACATGAGTGA 2559
Db 2521 AATGACGGGAACACAGGAATCAGGGGGAACATGAGTGA 2559

RESULT 2

US-10-035-045-3
; Sequence 3, Application US/10035045
; Publication No. US2003005448A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIAODONG
; APPLICANT: STASZEWSKI, LENA
; APPLICANT: O'CONNELL, SHAWN
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: TLR TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003-0280681
; CURRENT APPLICATION NUMBER: US/10/035,045
; PRIOR FILING DATE: 2002-01-03
; PRIOR APPLICATION NUMBER: 60/259,227
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: 60/284,547
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-035-045-3

Query Match 100.0%; Score 2559; DB 9; Length 2559;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 1 AAGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
QY 61 GCCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
Db 61 GCCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 121 TTCCCTGCTGGGAGGCGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 121 TTCCCTGCTGGGAGGCGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 181 GTGTGACACAGTGTCTCTCAACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 181 GTGTGACACAGTGTCTCTCAACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 241 GAGGAGATCAACACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 241 GAGGAGATCAACACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 301 GATACGCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 301 GATACGCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360

QY 361 GGCAGCCGCGAATGCGCGGCTACTGCAACTACAGAGTACAGAGCCCGCTGCTGCTGCT 420
Db 361 GGCAGCCGCGAATGCGCGGCTACTGCAACTACAGAGTACAGAGCCCGCTGCTGCTGCTGCT 420
QY 421 GTGATCGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 421 GTGATCGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 CTCATGCCCCAGGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 481 CTCATGCCCCAGGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 CCTCTCTTCTTCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db 541 CCTCTCTTCTTCCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
QY 601 CTGCAAGAGTTGCGGTGGAATGAGTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 CTGCAAGAGTTGCGGTGGAATGAGTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 CAGGGCTGAGCATCTTCTGCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 661 CAGGGCTGAGCATCTTCTGCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 GGGCTGGTGGGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
Db 721 GGGCTGGTGGGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 781 CACAGGTGAAACAGAGCAGCTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 781 CACAGGTGAAACAGAGCAGCTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 841 CAGGCGCTTCTTCACTACAGCATCAGCATCAGCATCAGCATCAGCATCAGCATCAGCATCAG 900
Db 841 CAGGCGCTTCTTCACTACAGCATCAGCATCAGCATCAGCATCAGCATCAGCATCAGCATCAG 900
QY 901 GAGGCGTGGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 GAGGCGTGGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 961 GTGCTTGGCTTCTCCTCAGAGGGGCTGCCAGCTGCAAGATTCCTCCAGTACGTGAAGAG 1020
Db 961 GTGCTTGGCTTCTCCTCAGAGGGGCTGCCAGCTGCAAGATTCCTCCAGTACGTGAAGAG 1020
QY 1021 CACCTGGGCTTGGCCACGACCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1021 CACCTGGGCTTGGCCACGACCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1081 CTGGAGAGAGAGTGGTGGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 1081 CTGGAGAGAGAGTGGTGGGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
QY 1141 GTGAGGCGAGGGCTAAATACCAACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
Db 1141 GTGAGGCGAGGGCTAAATACCAACAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
QY 1201 GCCCAGGCGCTGCAACACTTTCAGTGAACGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTG 1260
Db 1201 GCCCAGGCGCTGCAACACTTTCAGTGAACGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTG 1260
QY 1261 GTGAAGCCCTGGCAGTCTGCTGAGAAATGTAACCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db 1261 GTGAAGCCCTGGCAGTCTGCTGAGAAATGTAACCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
QY 1321 CGGCTGCGGTGAGACAGCAGCGGAAATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1380
Db 1321 CGGCTGCGGTGAGACAGCAGCGGAAATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1380
QY 1381 TGGCAGGGCTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
Db 1381 TGGCAGGGCTGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440

QY	1441	GAGGCGCTGAAAGATCCGCTGGGACACGTGTCGAAACACAGAAACCCGCTGTCCCGGGCTCG	1500
Db	1441	GAGGCGCTGAAAGATCCGCTGGGACACGTGTCGAAACACAGAAACCCGCTGTCCCGGGCTCG	1500
QY	1501	CGGAGATGCCAGAGAGGGGCGACAGTGCACCAGGATTCACATCCCTGCTGTACAC	1560
Db	1501	CGGAGATGCCAGAGAGGGGCGACAGTGCACCAGGATTCACATCCCTGCTGTACAC	1560
QY	1561	TGTGTGACATGCGAGGCGGGGCACTACCGGCAAAAACCCAGAGACATGCGCTGCACCTTT	1620
Db	1561	TGTGTGACATGCGAGGCGGGGCACTACCGGCAAAAACCCAGAGACATGCGCTGCACCTTT	1620
QY	1621	TGTGGCCAGGATAGTGTGTCCCGGAGCGCAACACACGCTGTTCCCGCGGAGGTCTCG	1680
Db	1621	TGTGGCCAGGATAGTGTGTCCCGGAGCGCAACACACGCTGTTCCCGCGGAGGTCTCG	1680
QY	1681	TTCTGTGGCATAGGGGCGAGCGGCTGTAGCTGTGCTGCTCTCTCTGATAGCTGGAGCTG	1740
Db	1681	TTCTGTGGCATAGGGGCGAGCGGCTGTAGCTGTGCTGCTCTCTGATAGCTGGAGCTG	1740
QY	1741	GGCCTTGTGCTGGCTGCTTTGGGGGCTGTTCGTTACCATCGGAGACCCCACTGGTTACG	1800
Db	1741	GGCCTTGTGCTGGCTGCTTTGGGGGCTGTTCGTTACCATCGGAGACCCCACTGGTTACG	1800
QY	1801	GGCTCGGGGGGGGGGGGGGGGCGCTGTTGGCCGTTGGTGTGCTGTGGGCGTGGTCTGCTACG	1860
Db	1801	GGCTCGGGGGGGGGGGGGGGGCGCTGTTGGCCGTTGGTGTGCTGTGGGCGTGGTCTGCTACG	1860
QY	1861	GTGCTCTGTTCCCTGTGGCGACGCCAGCCCTGCGCCGATGCTGTGGCCAGACGCTTGTCC	1920
Db	1861	GTGCTCTGTTCCCTGTGGCGACGCCAGCCCTGCGCCGATGCTGTGGCCAGACGCTTGTCC	1920
QY	1921	CACCTCCCGCTCACGGGCTGCTGTAGCACACTTTCCTGTGACAGGCGCGGAGATCTTGCTG	1980
Db	1921	CACCTCCCGCTCACGGGCTGCTGTAGCACACTTTCCTGTGACAGGCGCGGAGATCTTGCTG	1980
QY	1981	GAGTCACAACGCGCTCTGACCTGGGAGAGACCGGGTGAAGTGTGCTGTGGGGGGGCGCTCG	2040
Db	1981	GAGTCACAACGCGCTCTGACCTGGGAGAGACCGGGTGAAGTGTGCTGTGGGGGGGCGCTCG	2040
QY	2041	GCCCTGGGTGGTGTGCTGTGAGGCGCATGCTGTGTGAGAGTGTGCACTGTGACCTGGATCTG	2100
Db	2041	GCCCTGGGTGGTGTGCTGTGAGGCGCATGCTGTGTGAGAGTGTGCACTGTGACCTGGATCTG	2100
QY	2101	GTGGCCTTCCCGCGGAGGTGTGACGAGACTGGACATGCTGCCACGGAAGCGCTGTGGT	2160
Db	2101	GTGGCCTTCCCGCGGAGGTGTGACGAGACTGGACATGCTGCCACGGAAGCGCTGTGGT	2160
QY	2161	CACGTGGCGACACGCTCTGAGCTGGGCGACCTTCGCGGCGTAAAGCGCACACCAATGCCACGCTG	2220
Db	2161	CACGTGGCGACACGCTCTGAGCTGGGCGACCTTCGCGGCGTAAAGCGCACACCAATGCCACGCTG	2220
QY	2221	GCCCTTCTGCTGCTGCGGAGCACTTTCCTGTGTGTCGAGGACGACCGCGGCTGTGTACAACGCT	2280
Db	2221	GCCCTTCTGCTGCTGCGGAGCACTTTCCTGTGTGTCGAGGACGACCGCGGCTGTGTACAACGCT	2280
QY	2281	GGCCGCTGCGCTCACCTTGTGCATGTGCGCTACTTCATCACTCGTGGCTCTCTTTGTGGCC	2340
Db	2281	GGCCGCTGCGCTCACCTTGTGCATGTGCGCTACTTCATCACTCGTGGCTCTCTTTGTGGCC	2340
QY	2341	CTCTGTGGCAATGTGACGAGTGGTCTCAGGGCCCCGCGTGTGCAATGTGGGCGCCTCTGTCT	2400
Db	2341	CTCTGTGGCAATGTGACGAGTGGTCTCAGGGCCCCGCGTGTGCAATGTGGGCGCCTCTGTCT	2400
QY	2401	TGTGTCTGGGACATCGTGGCTGCTTCACACTGTGCCAGGTGTATCTGTCTATCGGGAG	2460
Db	2401	TGTGTCTGGGACATCGTGGCTGCTTCACACTGTGCCAGGTGTATCTGTCTATCGGGAG	2460
QY	2461	CGAAGGCTCAACACCCCGAGTCTTCTCTGTGAGAGGGGCGCGGGAGATGCCCAAGGCGAG	2520
Db	2461	CGAAGGCTCAACACCCCGAGTCTTCTCTGTGAGAGGGGCGCGGGAGATGCCCAAGGCGAG	2520
QY	2521	AATGACGGGAACACAGAAATCAGGGGAACATGATGA	2559

[illegible]

QY	661	CAGGCGCTGAGCAACATTCTTCTGGGCCCTTGGCCCGGCGGACAGCGGGCATCTGCATCGCCAGCAG	720
Db	661	CAGGGCTCGACACATTCTTCTGGGCCCTTGGCCCGGCGGACAGCGGGATCTGCATCCCGCAGCAG	720
QY	721	GCGCTGGTGGCGCTGGCCCGCTGGCGAATGATCTCGGCGCTGGGAAAGTGCAGACCTCTGT	780
Db	721	GCGCGTGGTGGCGCTGGCCCGCTGGCGAATGATCTCGGCGCTGGGAAAGTGCAGACACTCTGT	780
QY	781	CACCAAGTGAACCAAGACAGCGTGCAGAGTGTGTCTGTCTTTCGCTCCGTGCACGCCGCC	840
Db	781	CACCAAGTGAACCAAGACAGCGTGCAGAGTGTGTCTGTCTTTCGCTCCGTGCACGCCGCC	840
QY	841	CACGCGCCCTTCAACTACATACATACAGAGAGGCTCTGCCCCAABAGTGGGGTGGCCAGC	900
Db	841	CACGCGCCCTTCAACTACATACATACAGAGAGGCTCTGCCCCAABAGTGGGGTGGCCAGC	900
QY	901	GAGGCGCTGGCTGACCTTCTGACCTGTCTATGGGGCTCCCGGCATGGCCAGTGGGCGACG	960
Db	901	GAGGCGCTGGCTGACCTTCTGACCTGTCTATGGGGCTCCCGGCATGGCCAGTGGGCGACG	960
QY	961	GTCCTTGGCTTCTCTCCAGAGGGGTGCCCACTGTACAGATTTCCCAAGTACGTGAACG	1020
Db	961	GTCCTTGGCTTCTCTCCAGAGGGGTGCCCACTGTACAGATTTCCCAAGTACGTGAACG	1020
QY	1021	CACCTGGCCCTTGGGCGACCGACCCGGCCCTTGTGCTGGCCCTGGGGGAGAGGGAGCAGGT	1080
Db	1021	CACCTGGCCCTTGGGCGACCGACCCGGCCCTTGTGCTGGCCCTTGGGGGAGAGGGAGCAGGT	1080
QY	1081	CTGAGAGAGAGACGTGTGGGCGACGGCTGCCCGACAGTGTACATGATCAGCTGCAGAC	1140
Db	1081	CTGAGAGAGAGACGTGTGGGCGACGGCTGCCCGACAGTGTACATGATCAGCTGCAGAC	1140
QY	1141	GTCAGCGAGAGGCTAAATTCACACACAGAGCTTCTGTGTACGACGTGTATATAGCGTG	1200
Db	1141	GTCAGCGAGAGGCTAAATTCACACACAGAGCTTCTGTGTACGACGTGTATATAGCGTG	1200
QY	1201	GCCAGGCGCTTGCAACAACACTTTCATGTGCACAGCTCAGGCTGCCCGCGCAGGACCCC	1260
Db	1201	GCCAGGCGCTTGCAACAACACTTTCATGTGCACAGCTCAGGCTGCCCGCGCAGGACCCC	1260
QY	1261	GTCGAAGCCCTGGGCGCTCGTGGAGAACATGTAAACCTTACCTTCCACAGTGGGCGGGTG	1320
Db	1261	GTCGAAGCCCTGGGCGCTCGTGGAGAACATGTAAACCTTACCTTCCACAGTGGGCGGGTG	1320
QY	1321	CCGCTGCGGTTTCGACACACACCGGAAACGTGTGACATGAGATAGACCTGAAGCTGTGGGTG	1380
Db	1321	CCGCTGCGGTTTCGACACACACCGGAAACGTGTGAGATAGAGACTGAAGCTGTGGGTG	1380
QY	1381	TGGCAGGGCTCAGTGTGCCCAAGGCTTCACAGACGTGTGGATTTCAAGGCGACTCAGGACA	1440
Db	1381	TGGCAGGGCTCAGTGTGCCCAAGGCTTCACAGACGTGTGGATTTCAAGGCGACTCAGGACA	1440
QY	1441	GAGCGCCGGAAGATCCGCTGGACACAGTCTGAAACAACAABAACCCGTGCCGGGGTCGCG	1500
Db	1441	GAGCGCCGGAAGATCCGCTGGACACAGTCTGAAACAACAABAACCCGTGCCGGGGTCGCG	1500
QY	1501	CGCGAGTGTCCAGGAGGGGCGCAGGTGTCCCGGGGTTCMAAGGGTTCCACTCCGTGTACAC	1560
Db	1501	CGCGAGTGTCCAGGAGGGGCGCAGGTGTCCCGGGGTTCMAAGGGTTCCACTCCGTGTACAC	1560
QY	1561	TGTGTGGACCTCGAGGGGGGCGGACCTTACCGGCAAAACCCAGACAGCATGGCTCGACCTTT	1620
Db	1561	TGTGTGGACCTCGAGGGGGGCGGACCTTACCGGCAAAACCCAGACAGCATGGCTCGACCTTT	1620
QY	1621	TGTGGCAGGATGAGTGTCCCGGAGCGAAGACACAGCTGTCTCCCGCGCAGGTCTCGG	1680
Db	1621	TGTGGCAGGATGAGTGTCCCGGAGCGAAGACACAGCTGTCTCCCGCGCAGGTCTCGG	1680
QY	1681	TTCTCTGGCATAGGGGCGAGCGCGGTGTGGTGTCTGTCTCTCTCTAGGCTTGGCGGTG	1740
Db	1681	TTCTCTGGCATAGGGGCGAGCGCGGTGTGGTGTCTGTCTCTCTCTAGGCTTGGCGGTG	1740

QY	1741	GGCTTGTGCTGAGTGGCTTTTGGGGCTGTTCCTTACACCAATCGGGAGACGCCACTGGTTAG	1800
Db	1741	GGCTTTGTGCTGAGTGGCTTTTGGGGCTGTTCCTTACACCAATCGGGAGACGCCACTGGTTAG	1800
QY	1801	GCCTCGGGGGGGGGCCCTTGGGCTGCTTTTGGGCGTGGTGTGGCTGGGCTGGTGTGGCTTACG	1860
Db	1801	GCCTCGGGGGGGGGCCCTTGGGCTGCTTTTGGGCGTGGTGTGGCTGGGCTGGTGTGGCTTACG	1860
QY	1861	GTCCCTCCTGTTCCTTGGGCGAGCCAGCCCTGTGCCGATGCTGTGGCCAGAGAGGCCCTTGTCC	1920
Db	1861	GTCCCTCCTGTTCCTTGGGCGAGCCAGCCCTGTGCCGATGCTGTGGCCAGAGAGGCCCTTGTCC	1920
QY	1921	CACCTCCCGCTCACGGGCTGCTGTAGCACACTTCTCTCAGAGCGGCGAGATCTTGTGTG	1980
Db	1921	CACCTCCCGCTCACGGGCTGCTGTAGCACACTTCTCTCAGAGCGGCGAGATCTTGTGTG	1980
QY	1981	GAGTCAGAACTGCTCTGAGCTGTGGGAGAGCCGCGCTGAGTGGCTGTGGGCGGGGGCCCTGG	2040
Db	1981	GAGTCAGAACTGCTCTGAGCTGTGGGAGAGCCGCGCTGAGTGGCTGTGGGCGGGGGCCCTGG	2040
QY	2041	GCCTGGCGTGGTGGTGTGCTGTGGCCATGCTGGTGGAGAGTGGCATGTGACCTGTGATCCTG	2100
Db	2041	GCCTGGCGTGGTGGTGTGCTGTGGCCATGCTGGTGGAGAGTGGCATGTGACCTGTGATCCTG	2100
QY	2101	GTGGCTTTCCTCCGCGGAGGTGTGTAGCGAGCTGCGACATGCTGTCCACAGGAGGCGCTGTG	2160
Db	2101	GTGGCTTTCCTCCGCGGAGGTGTGTAGCGAGCTGCGACATGCTGTCCACAGGAGGCGCTGTG	2160
QY	2161	CACATCGCGACACGCTCTCTGGGTACGTTTGGGCTTAGCGCAGCGCACCAATATCCACGCTG	2220
Db	2161	CACATCGCGACACGCTCTCTGGGTACGTTTGGGCTTAGCGCAGCGCACCAATATCCACGCTG	2220
QY	2221	GCCTTTCCTGCTCTGCTGTGGGCACTTTCCTGGTGGAGAGCGAGCGCGGCGTGTACAAACGT	2280
Db	2221	GCCTTTCCTGCTCTGCTGTGGGCACTTTCCTGGTGGAGAGCGAGCGCGGCGTGTACAAACGT	2280
QY	2281	GCCCGTGGCCTCACCTTGTGGCATGTGGGCTTACTTATCATCATCGTGGGTCTCTTGTGCC	2340
Db	2281	GCCCGTGGCCTCACCTTGTGGCATGTGGGCTTACTTATCATCATCGTGGGTCTCTTGTGCC	2340
QY	2341	CTCCGCGGCATGTGTAGAGTGTGTCTCAGGCGCGCGCTGACAAATGGGGCCCTCTGTGCTC	2400
Db	2341	CTCCGCGGCATGTGTAGAGTGTGTCTCAGGCGCGCGCTGACAAATGGGGCCCTCTGTGCTC	2400
QY	2401	TGTGTCTGGGCACTGTGGTGTGCTTCCACTGTCCAGGTGTACTGTCTATGTGGGAG	2460
Db	2401	TGTGTCTGGGCACTGTGGTGTGCTTCCACTGTCCAGGTGTACTGTCTATGTGGGAG	2460
QY	2461	CCAGGGCTCAACACCCCCGAGTCTTCTGTGGAGGGGGCCCTGGGGATGCCCCAGAGCCAG	2520
Db	2461	CCAGGGCTCAACACCCCCGAGTCTTCTGTGGAGGGGGCCCTGGGGATGCCCCAGAGCCAG	2520
QY	2521	AATGACGGGAACACAGGAATTCAGGGGAAACATGATGA	2559
Db	2521	AATGACGGGAACACAGGAATTCAGGGGAAACATGATGA	2559
RESULT 4			
US-09-897-427A-7			
; Sequence 7, Application US/09897427A			
; Patent No. US20020160424A1			
; GENERAL INFORMATION:			
; APPLICANT: ADLER, JON ELLIOT			
; APPLICANT: LI, XIADONG			
; APPLICANT: STRZEMSKI, LENA			
; APPLICANT: XU, HONG			
; APPLICANT: EHEVERRI, FERNANDO			
; TITLE OF INVENTION: T1R HETERO-OLIGOMERIC TASTE RECEPTORS			
; FILE REFERENCE: 078003-0282558			
; CURRENT APPLICATION NUMBER: US/09/897,427A			
; CURRENT FILING DATE: 2001-07-03			
; NUMBER OF SEQ ID NOS: 10			
; SOFTWARE: PatentIn Ver. 2.1			

```

; SEQ ID NO 7
; LENGTH: 2559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-427A-7

```

Query Match	99.9%	Score 2557.4	DB 9	Length 2559
Best Local Similarity	100.0%	Pred.No. 0		
Matches 2558	Conservative 0	Mismatches 1	Indels 0	Gaps 0

OY	1	ATGCGTGGGACCCGTGGTGCCTCTGAGGGGCTGACGCTCTGGGGCTGTCGCGACGCTGGGAGACGGG	60
Db	1	ATGCTGGGACCCCTGCTGTCTCTTGGGCTCAAGCTCTGGGGCTCTCCGACGCTGGGAGACGGG	60
OY	61	GCCCCATTTGTGCTGTCAACAGCACTTAGATGAAAGGGAGACTACGTCTGGGGGGGGCTG	120
Db	61	GCCCCATTTGTGCTGTCAACAGCACTTAGATGAAAGGGAGACTACGTCTGGGGGGGGCTG	120
OY	121	TTCCCCCTGGGGCGAGGCCGAGAGAGGCTTGGCTCCGCGACCGGGACAGGGCCGACAGCCCT	180
Db	121	TTCCCCCTGGGGCGAGGGCCGAGAGAGGCTGGCTCCGCGACCGGGACAGGGCCGACAGCCCT	180
OY	181	GTGTCAACAGAGTTTCCTCTCAAGAGGCGCTGTCTGGGACACTGGGCGCATGAATAATGGCCGTG	240
Db	181	GTGTCAACAGAGTTTCCTCTCAAGAGGCGCTGTCTGGGACACTGGGCGCATGAATAATGGCCGTG	240
OY	241	GAGGAGATCAACACACAGTGGATCTGTGTGCCCGGGCTGGCGCTGGGCTACGACCTCTTT	300
Db	241	GAGGAGATCAACACACAGTGGATCTGTGTGCCCGGGCTGGCGCTGGGCTACGACCTCTTT	300
OY	301	GATACGTGCTGGAGCCCTGTGTGTGGCCATGAAAGCCAGGCTCATGTTCTTGGCCAGGCA	360
Db	301	GATACGTGCTGGAGCCCTGTGTGTGGCCATGAAAGCCAGGCTCATGTTCTTGGCCAGGCA	360
OY	361	GGCAGCCCGCAGATCGCCGCTACTGCAATACACAGCATACCGAGTACCGAGCCCGGCTGGCT	420
Db	361	GGCAGCCCGCAGATCGCCGCTACTGCAATACACAGCATACCGAGTACCGAGCCCGGCTGGCT	420
OY	421	GTCAATCGGGCCCACTGTCTCAGAGCTCGCCATGGTCTCAACCGGCAAGTTCTTCAGCTTCTTC	480
Db	421	GTCAATCGGGCCCACTGTCTCAGAGCTCGCCATGGTCTCAACCGGCAAGTTCTTCAGCTTCTTC	480
OY	481	CTCATGCCCCCAGGTACGTACGAGGTCTACATGGAAGTGTGTGAGCGCCCGGAGACCTTC	540
Db	481	CTCATGCCCCCAGGTACGTACGAGGTCTACATGGAAGTGTGTGAGCGCCCGGAGACCTTC	540
OY	541	CCCTCTCTTTCGAGCCGTCGCCAGGAGCCGTGTGACGTGAGAGGCGCGCGCGGAGACTG	600
Db	541	CCCTCTCTTTCGAGCCGTCGCCAGGAGCCGTGTGACGTGAGAGGCGCGCGCGGAGACTG	600
OY	601	CTCGAGAGTTCTGGCTGGAATCTGGGTGGCCCTCTGGGAGAGCGACGAGATACGGCCGG	660
Db	601	CTCGAGAGTTCTGGCTGGAATCTGGGTGGCCCTCTGGGAGAGCGACGAGATACGGGTGGCCGG	660
OY	661	CAGGGCTTGAGCATCTCTCGGGCCCTGGCCGCGGACGCGGCACTCTGATTCGCGACAG	720
Db	661	CAGGGCTTGAGCATCTCTCGGGCCCTGGCCGCGGACGCGGCACTCTGATTCGCGACAG	720
OY	721	GGCCGTGGGCGCGGCCCGGTGCGGATGACATCGGCGCTGGGGAAGGTGACAGAGTCTCTG	780
Db	721	GGCCGTGGGCGCGGCCCGGTGCGGATGACATCGGCGCTGGGGAAGGTGACAGAGTCTCTG	780
OY	781	CACCAAGTGAACACAGAGAGCGGTGAGGTGTGTCTGTTCGCTCCGCTCCGATGACGCGCC	840
Db	781	CACCAAGTGAACACAGAGAGCGGTGAGGTGTGTCTGTTCGCTCCGCTCCGATGACGCGCC	840
OY	841	CACGCCCTTTCACACTACAGCATCAGACGAGGCTCTCGCCCAAGGTGTGGGTGGCCAGC	900
Db	841	CACGCCCTTTCACACTACAGCATCAGACGAGGCTCTCGCCCAAGGTGTGGGTGGCCAGC	900
OY	901	GAGGCTCTGGGCTGACTCTGACACTGTGTATGGGGGTGCCCGGCAATGGCCAGATGGGCAG	960
Db	901	GAGGCTCTGGGCTGACTCTGACACTGTGTATGGGGGTGCCCGGCAATGGCCAGATGGGCAG	960

[illegible]

[illegible]

RESULT 5
US-10-261-482-1

```

: Sequence 1 Application US/10261482
: Publication No. US20030036089A1
:
: GENERAL INFORMATION:
:
: APPLICANT: WET, Ming-Hui et al
:
: TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
:
: TITLE OF INVENTION: RECEPTORS, NOCLEIC ACID MOLECULES ENCODING HUMAN GPCR
:
: TITLE OF INVENTION: PROTEINS, AND USES THEREOF
:
: FILE REFERENCE: C0000869CON
:
: CURRENT APPLICATION NUMBER: US/10/261,482
:
: CURRENT FILING DATE: 2002-10-02
:
: PRIOR APPLICATION NUMBER: 09/684,393
:
: PRIOR FILING DATE: 2000-10-10
:
: PRIOR APPLICATION NUMBER: 60/172,600
:
: PRIOR FILING DATE: 1999-12-20
:
: NUMBER OF SEQ ID NOS: 10
:
: SOFTWARE: FASTSEQ for Windows Version 4.0
:
: SEQ ID NO 1
:
: LENGTH: 3458
:
: TYPE: DNA
:
: ORGANISM: HUMAN
:
: US-10-261-482-1

```

Query Match	99.9%	Score 2557.4;	DB 9;	Length 3458;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2558; Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;

0y 1 ATCTGGGGCCCTGCTCTCTCTGGGCTCAGCCTCTGGGCTCTCTCTGCACCTCTGGACGGG 60
0y 30 ATCTGGGGCCCTGCTCTCTCTGGGCTCAGCCTCTGGGCTCTCTCTGCACCTCTGGACGGG 89
0y 61 GCCCATTTGTGCTGTCTCAGCAGCAACTTAGATGAGAGGGGGGACTGCTGTGGGGGGCTG 120
Db 90 GCCCATTTGTGCTGTCTCAGCAGCAACTTAGATGAGAGGGGGGACTGCTGTGGGGGGCTG 149

QY	121	TTCCCCCTGGGCGAGGCCGAGAGGCTGGCCTCCGACGCCGACAGGCCACGACGCTT	180
Db	150	TTCCCCCTGGGCGAGGCCGAGAGGCGCTGGCCTCCGAGGCCGACAGGCCACGAGGCCCT	209
QY	181	GTGTGCACACAGGTTCTCTTAAACGGCGCTGTGGGCACTGGCCATGAAATGGACGCTG	240
Db	210	GTGTGCACACAGGTTCTCTTAAACGGCGCTGTGGGCACTGGCCATGAAATGGACGCTG	269
QY	241	GAGGAGATCAACAACAAGTGGATCTGTCTGTCCCGGGCTCGCGCTGGGTACGACCTTTT	300
Db	270	GAGGAGATCAACAACAAGTGGATCTGTCTGTCCCGGGCTCGCGCTGGGTACGACCTTTT	329
QY	301	GATACGTCTCGAGAGCTGTGTGTGGCCATGAAAGCCAGCCTATGTCTTGGCCAAAGCA	360
Db	330	GATACGTCTCGAGAGCTGTGTGTGGCCATGAAAGCCAGCCTATGTCTTGGCCAAAGCA	389
QY	361	GGCAGCGCGACATCGCGCGCTACTGCACTACACATACACATACACCCCGTGTGCTGCT	420
Db	390	GGCAGCGCGACATCGCGCGCTACTGCACTACACATACACATACACCCCGTGTGCTGCT	449
QY	421	GTCAATCGGGCCCCACTCGTACAGACTGCGCCATGTGTAACCGCAAGTTCTTACGTTCTT	480
Db	450	GTCAATCGGGCCCCACTCGTACAGACTGCGCCATGTGTAACCGCAAGTTCTTACGTTCTT	509
QY	481	CTCATGCCCCAGGTCACTACGGTGTAGCATGGAAGCTCTAGAGGCCGGAGACACTTC	540
Db	510	CTCATGCCCCAGGTCACTACGGTGTAGCATGGAAGCTCTAGAGGCCGGAGACACTTC	569
QY	541	CCCTCTCTTCTTCGCGACCGTGGCCAGCGACTGTGTACAGCTGAAGCGCGCGCGGACGTG	600
Db	570	CCCTCTCTTCTTCGCGACCGTGGCCAGCGACTGTGTACAGCTGAAGCGCGCGCGGACGTG	629
QY	601	CTGCGAGAGTTCCGAGCTGGAATGGGTGGCGCGCTGGGAGGAGGACAGATGACGGCGG	660
Db	630	CTGCGAGAGTTCCGAGCTGGAATGGGTGGCGCGCTGGGAGGAGGACAGATGACGGCGG	689
QY	661	CAGGCGCTGAGCATTTCTCGGCCCTGTGGCCGGCAGCGCGGACATCTGCATCGCGCAGAG	720
Db	690	CAGGCGCTGAGCATTTCTCGGCCCTGTGGCCGGCAGCGCGGACATCTGCATCGCGCAGAG	749
QY	721	GGCCGTGGCGCGCGCGCGCGCGCGCGCATGATCGCGCGCTGGGAAAGTGCAGGACGCTCTG	780
Db	750	GGCCGTGGCGCGCGCGCGCGCGCGCATGATCGCGCGCTGGGAAAGTGCAGGACGCTCTG	809
QY	781	CACCAAGGTGAACAGACAGCGAGTGTGAGTGTCTCTGTGCGCTCCGTGCAGCCGCC	840
Db	810	CACCAAGGTGAACAGACAGCGAGTGTGAGTGTCTCTGTGCGCTCCGTGCAGCCGCC	869
QY	841	CACGCCCTTTCAACTACAGATACAGACAGAGGCTTCGCGCCAAAGTGTGGTGGCCACG	900
Db	870	CACGCCCTTTCAACTACAGATACAGACAGAGGCTTCGCGCCAAAGTGTGGTGGCCACG	929
QY	901	GAGGCGCTGGCTGACCTGTGACTGTGATGGGGCTCCCGGATGGCCAGATGGGACGACG	960
Db	930	GAGGCGCTGGCTGACCTGTGACTGTGATGGGGCTCCCGGATGGCCAGATGGGACGAG	989
QY	961	GTGCTTGGCTTCTCCAGAGGGGTGCCACGCTGACAGATTTCCCAAGTACGTGAAGAC	1020
Db	990	GTGCTTGGCTTCTCCAGAGGGGTGCCACGCTGACAGATTTCCCAAGTACGTGAAGAC	1049
QY	1021	CACCTGGGCGCTGGGACACGCGCGCTTGTGTCTGGCCCTGGGAGAGAGGAGAGAGT	1080
Db	1050	CACCTGGGCGCTGGGACACGCGCGCTTGTGTCTGGCCCTGGGAGAGAGGAGAGAGT	1109
QY	1081	CTGAGAGAGGAGCTGTGTGGGCCAGCGCTGCGCGAGTGTACTGCATCACGCTGCAGAAC	1140
Db	1110	CTGAGAGAGGAGCTGTGTGGGCCAGCGCTGCGCGAGTGTACTGCATCACGCTGCAGAAC	1169
QY	1141	GTGAGCGAGGGGTAAATTCACACACAGACTTCTGTGTCTAGCGAGCTGTGATATGCGTG	1200
Db	1170	GTGAGCGAGGGGTAAATTCACACACAGACTTCTGTGTCTAGCGAGCTGTGATATGCGTG	1229
QY	1201	GCCCAAGGCCCTGCAACAACACTTTCATGTGACGCTTACAGGCTTACGCCCGCGGACGACCC	1260

Db 1230 GCGGAGGCGCTGACAAACCTTTCAGTGACAGCGCTCAGGCTGCCCGCAGAGACCCC 1289
 QY 1261 GTGAAGCCCTGGAGCTCTGGAGAACATGTACAACTGACCTTCCAGTGGGGGGCTG 1320
 Db 1290 GTGAAGCCCTGGAGCTCTGGAGAACATGTACAACTGACCTTCCAGTGGGGGGCTG 1349
 QY 1321 CCAGTGGGGTTCAGACAGCAGCGGAAACGTGACATGAGTACGACCTGAGCTGGGGT 1380
 Db 1350 CCGCTGGCGTTCAGACAGCAGCGGAAACGTGACATGAGTACGACCTGAGCTGGGGT 1409
 QY 1381 TGGAGGGGCTAGTGGCCAGGCTCCAGACGTGGGAGTTCACAGCGAGGCTCAGAGCA 1440
 Db 1410 TGGAGGGGCTAGTGGCCAGGCTCCAGACGTGGGAGTTCACAGCGAGGCTCAGAGCA 1469
 QY 1441 GAGCGCTGAAAGATCCGCTGGCACACGTCTGACAAACAGAAAGCCGCTTCCGGTCTCG 1500
 Db 1470 GAGGCGCTGAAGATCCGCTGGCACACGTCTGACAAACAGAAAGCCGCTTCCGGTCTCG 1529
 QY 1501 CGGCAGTGCCAGAGAGGCGCAGGTGCGCGGGTCAAGGGGGTCCACTCTGCTAGAC 1560
 Db 1530 CGGCAGTGCCAGAGAGGCGCAGGTGCGCGGGTCAAGGGGGTCCACTCTGCTAGAC 1589
 QY 1561 TGTGTGAGCTGCGAGGGGAGCTACCGGAAACCCAGACGACATGCGCTGACCTT 1620
 Db 1590 TGTGTGAGCTGCGAGGGGAGCTACCGGAAACCCAGACGACATGCGCTGACCTT 1649
 QY 1621 TGTGTGAGCTGCGAGGGGAGCTACCGGAGGAGACACAGCTCTCCGCGGAGGCTCGG 1680
 Db 1650 TGTGTGAGCTGCGAGGGGAGCTACCGGAGGAGACACAGCTCTCCGCGGAGGCTCGG 1709
 QY 1681 TTCTGTGAGTGGGCGAGCGCGGTGTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1740
 Db 1710 TTCTGTGAGTGGGCGAGCGCGGTGTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1769
 QY 1741 GGCTTGT 1800
 Db 1770 GGCTTGT 1829
 QY 1801 GGCTTGT 1860
 Db 1830 GGCTTGT 1889
 QY 1861 GTCTTGT 1920
 Db 1890 GTCTTGT 1949
 QY 1921 CACCTTCCGCTCAGGGGCTGT 1980
 Db 1950 CACCTTCCGCTCAGGGGCTGT 2009
 QY 1981 GAGTCAGAACTGT 2040
 Db 2010 GAGTCAGAACTGT 2069
 QY 2041 GCGT 2100
 Db 2070 GCGT 2129
 QY 2101 GT 2160
 Db 2130 GT 2189
 QY 2161 CACTGT 2220
 Db 2190 CACTGT 2249
 QY 2221 GCGT 2280
 Db 2250 GCGT 2309
 QY 2281 GCGGTTGGGCTCAGCTTGT 2340

Db 2310 GCGGTTGGGCTCAGCTTGT 2369
 QY 2341 CTCTGT 2400
 Db 2370 CTCTGT 2429
 QY 2401 TGT 2460
 Db 2430 TGT 2489
 QY 2461 CCAGGGCTCAACACCCCGAGTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2520
 Db 2490 CCAGGGCTCAACACCCCGAGTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2549
 QY 2521 AATGACGGGAAACAGGAAATCAGGGGAAACATGATGA 2559
 Db 2550 AATGACGGGAAACAGGAAATCAGGGGAAACATGATGA 2588

RESULT 6
 US-10-282-837-15
 ; Sequence 15, Application US/10282837
 ; Publication No. US20030082738A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc.
 ; TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
 ; FILE REFERENCE: 10448-020001
 ; CURRENT APPLICATION NUMBER: US/10/282,837
 ; CURRENT FILING DATE: 2002-10-29
 ; PRIOR APPLICATION NUMBER: US/09/796,338
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: US 60/186,059
 ; PRIOR FILING DATE: 2000-02-29
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 15
 ; LENGTH: 2559
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-282-837-15

Query Match 99.9%; Score 2555.8; DB 9; Length 2559;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGTGTGGGCGCTGT 60
 Db 1 ATGTGTGGGCGCTGT 60
 QY 61 GCCCATTTGT 120
 Db 61 GCCCATTTGT 120
 QY 121 TTCCCTGT 180
 Db 121 TTCCCTGT 180
 QY 181 GT 240
 Db 181 GT 240
 QY 241 GAGGAGATCAACAAGATGT 300
 Db 241 GAGGAGATCAACAAGATGT 300
 QY 301 GATACGT 360
 Db 301 GATACGT 360
 QY 361 GGCAGCCGCGACATGCGCGGCTACTGTACATACAGAGTACAGAGCCCGGTGTGTGTGTGT 420
 Db 361 GGCAGCCGCGACATGCGCGGCTACTGTACATACAGAGTACAGAGCCCGGTGTGTGTGTGT 420

421 GTCATCGGGCCCACTCGTCAGAGGCTGGCCATGTCACCGGCAAGTTCTTCAGCTTCTTC
421 GTGATCGGGCCCACTCGTCAGAGGCTGGCCATGTCACCGGCAAGTTCTTCAGCTTCTTC
481 CTGATGCCCCAGGTCAGTACGAGTGTCTAGCATGAGACTGCTGAGCGCCCCGGAGACCTTC
481 CTGATGCCCCAGGTCAGTACGAGTGTCTAGCATGAGACTGCTGAGCGCCCCGGAGACCTTC
541 CCTCTCTTCTTCGCGACCGTGGCCAGCGACCGTGTGCACTGACGCGCCCCCGGAGCTG
541 CCTCTCTTCTTCGCGACCGTGGCCAGCGACCGTGTGCACTGACGCGCCCCCGGAGCTG
601 CTGCAAGAGTTGGGCTGGAACTGGGAGGCGGCTGGGAGGAGAGAGATGACGCGCG
601 CTGCAAGAGTTGGGCTGGAACTGGGAGGCGGCTGGGAGGAGAGATGACGCGCG
661 CAGGGCTGTAGCATCTTCTCGGCGCTGGCGCGGACACGCGGATCTGATCGCGACGAG
661 CAGGGCTGTAGCATCTTCTCGGCGCTGGCGCGGACACGCGGATCTGATCGCGACGAG
721 GGCCTGTGGCCGCTGCGCCGCTGCGATGATGCTGGGCTGGGAGAGTGCAGACGCTG
721 GGCCTGTGGCCGCTGCGCCGCTGCGATGATGCTGGGCTGGGAGAGTGCAGACGCTG
781 CACCAAGTGAACAGAGCAGTGCAGAGTGTGCTGCTGCTGGCCCTGCGACGCGCGC
781 CACCAAGTGAACAGAGCAGTGCAGAGTGTGCTGCTGCTGGCCCTGCGACGCGCGC
841 CAGCGCTCTTCTACATACATACATACATACATACATACATACATACATACATACAT
841 CAGCGCTCTTCTACATACATACATACATACATACATACATACATACATACATACAT
901 GAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
901 GAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
961 GTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
961 GTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1021 CACTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1021 CACTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1081 CTGAG
1081 CTGAG
1141 GTGAG
1141 GTGAG
1141 GTGAG
1201 GCGGAG
1201 GCGGAG
1261 GTGAAG
1261 GTGAAG
1321 CCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1321 CCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1381 TGGCAG
1381 TGGCAG
1441 GAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1441 GAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

1501 CCGCAGTGTGAG
1501 CCGCAGTGTGAG
1561 TGTGTGAGTGTGAG
1561 TGTGTGAGTGTGAG
1621 TGTGTGAGTGTGAG
1621 TGTGTGAGTGTGAG
1681 TTCTGTGAGTGTGAG
1681 TTCTGTGAGTGTGAG
1741 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1741 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1801 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1801 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1861 GTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1861 GTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1921 CACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1921 CACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
1981 GAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG
1981 GAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG
2041 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2041 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2101 GTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2101 GTGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2161 CACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2161 CACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2221 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2221 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2281 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2281 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2341 GTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2341 GTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2401 TGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2401 TGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2461 CCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2461 CCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
2521 AATGAGGAG
2521 AATGAGGAG
2559 AATGAGGAG
2559 AATGAGGAG


```

Db 1801 GCGTCGGGGGGGGCCCTGGACCTCTTGGCTGTGTGCTGGGCTGTGCTGCTCAGC 1860
Oy 1861 GTCCCTCCCTTCCCTCCCTCCAGCCGACCCCTGCCCCGATGCTGGCCAGCAGCCCTGTGC 1920
Oy 1861 GTCCCTCCCTTCCCTCCCTCCAGCCGACCCCTGCCCCGATGCTGGCCAGCAGCCCTGTGC 1920
Oy 1921 CACCTCCCGCTCAGGGGCTGCTGAGCAGACATCTTCTCTCAGAGGCGCCGAGATCTTCTG 1980
Db 1921 CACCTCCCGCTCAGGGGCTGCTGAGCAGACATCTTCTCTCAGAGGCGCCGAGATCTTCTG 1980
Oy 1981 GAGTCAGAACCTGCTCTGAGCTGGGAGACACCGGCTGAGTGGCTGCTGCGGGGGCCCTG 2040
Db 1981 GAGTCAGAACCTGCTCTGAGCTGGGAGACACCGGCTGAGTGGCTGCTGCGGGGGCCCTG 2040
Oy 2041 GCGTCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2100
Db 2041 GCGTCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2100
Oy 2101 GTGGGCTTCCCGCGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2160
Db 2101 GTGGGCTTCCCGCGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2160
Oy 2161 CACTGCCGACACAGCTCTCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2220
Db 2161 CACTGCCGACACAGCTCTCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2220
Oy 2221 GCGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
Db 2221 GCGTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
Oy 2281 GCGCGTGGGCTCAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2340
Db 2281 GCGCGTGGGCTCAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2340
Oy 2341 CTCCTGGCCAATGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2400
Db 2341 CTCCTGGCCAATGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2400
Oy 2401 TGTGTCTGGGCTCAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
Db 2401 TGTGTCTGGGCTCAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
Oy 2461 CCAGGGCTCAGACCCCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2520
Db 2461 CCAGGGCTCAGACCCCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2520
Oy 2521 AATGACGGGAGACAGGAATCAGGGGAAACATGAGTGA 2559
Db 2521 AATGACGGGAGACAGGAATCAGGGGAAACATGAGTGA 2559

```

RESULT 8

```

; Sequence 13, Application US/10282837
; Publication No. US20030082738A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
; TITLE OF INVENTION: 52872, NOVEL G PROTEIN-COUPLED RECEPTORS AND USES THEREFOR
; FILE REFERENCE: 10448-020001
; CURRENT APPLICATION NUMBER: US/10/282,837
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US/09/796,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,059
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 3489
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(2607)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(3489)
; OTHER INFORMATION: n = A,T,C or G
US-10-282-837-13

```

Query Match 99.9%; Score 2555.8; DB 9; Length 3489;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Oy 1 ATGCTGGGGCCGACCTTCCGAGGAGCTCAGCCTCTGGGCTCTCTGCACTGAGCCAGGAG 60
Db 52 ATGCTGGGGCCGACCTTCCGAGGAGCTCAGCCTCTGGGCTCTCTGCACTGAGCCAGGAG 111
Oy 61 GCCCATTGTGCTGTACAGCACTTAAAGATGAAGGGGAGTACGTGTGGGGGGCTG 120
Db 112 GCCCATTGTGCTGTACAGCACTTAAAGATGAAGGGGAGTACGTGTGGGGGGGGCTG 171
Oy 121 TTCCCTGGGGGAGGCGGAGAGGCTGTGCTCCGAGCGGACAGGCGCCAGACCCCT 180
Db 172 TTCCCTGGGGGAGGCGGAGAGGCTGTGCTCCGAGCGGACAGGCGCCAGACCCCT 231
Oy 181 GTGTGACAGAGTCTCTCAAAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Db 232 GTGTGACAGAGTCTCTCAAAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 291
Oy 241 GAGGAGATCAACAACAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Db 292 GAGGAGATCAACAACAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 351
Oy 301 GATACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Db 352 GATACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 411
Oy 361 GGCAGCGGAGACATCCCGCTACTCACTACACAGCAGTACAGCCCGGTGTGTGTGT 420
Db 412 GGCAGCGGAGACATCCCGCTACTCACTACACAGCAGTACAGCCCGGTGTGTGTGTGT 471
Oy 421 GTTCATGGGCCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Db 472 GTTCATGGGCCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 531
Oy 481 CTCATCCCCAGGTGAGTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Db 532 CTCATCCCCAGGTGAGTACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 591
Oy 541 CCCTCTTCTTCCGACCGGTGCGCAGCAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Db 592 CCCTCTTCTTCCGACCGGTGCGCAGCAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 651
Oy 601 CTGACAGAGTGTGGGTGAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Db 652 CTGACAGAGTGTGGGTGAGAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 711
Oy 661 CAGGGGCTGAGATCTTCTGCGGCGGCGGCGGAGGAGGAGTGTGTGTGTGTGTGTGT 720
Db 712 CAGGGGCTGAGATCTTCTGCGGCGGCGGAGGAGGAGTGTGTGTGTGTGTGTGTGTGT 771
Oy 721 GGCCTGTGTGCGGCTGCGGCGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db 772 GGCCTGTGTGCGGCTGCGGCGGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 831
Oy 781 CACACAGTAAACAGACAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db 832 CACACAGTAAACAGACAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 891
Oy 841 CACGCGCTTCAACTACAGCATCAGCAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Db 892 CACGCGCTTCAACTACAGCATCAGCAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 951

```

QY	901	GAGCCTCGGCTGACCTCTGACCTGTCATGAGGGGCTCCCGGATGGCCCAATGATGGGACG	960
Db	992	GAGGCTGGGCTGACCTCTGACCTGTCATGAGGGGCTCCCGGATGGCCCAATGATGGGACG	10111
QY	961	GTGCTTGAGCTTCCTCCAGAGGGGTGCCAGACTGACAGAGTTCCCCAGTACTGTAAAGACG	10202
Db	1012	GTGCTTGAGCTTCCTCCAGAGGGGTGCCAGACTGACAGAGTTCCCCAGTACTGTAAAGACG	10711
QY	1021	CACCTGGGCTTCGGCCACCAGACCCGGGCTTGTGTCTGCCCCGTGGGGAGAGGAGACAGGAT	10800
Db	1072	CACCTGGGCTTCGGCCACCAGACCCGGGCTTGTGTCTGCCCCGTGGGGAGAGGAGACAGGAT	11311
QY	1081	CTGGAGGAGGACGTGTGGGACACAGCGTGGCCCCCATGTGACGTGATCAGCGTGCAGAC	11400
Db	1132	CTGGAGGAGGACGTGTGGGACACAGCGTGGCCCCCATGTGACGTGATCAGCGTGCAGAC	11911
QY	1141	GTGAGCGGAGGGCTAAATTCACACACAGAGTTCTGTCTACGCAAGCTGTATAGCGTG	12000
Db	1192	GTGAGCGGAGGGCTAAATTCACACACAGAGTTCTGTCTACGCAAGCTGTATAGCGTG	12511
QY	1201	GCCAGAGCCCTGGCAACACTCTTCAGTGAACGCCCTCAGCGTGCCCGCGAGAGACCC	12600
Db	1252	GCCAGAGCCCTGGCAACACTCTTCAGTGAACGCCCTCAGCGTGCCCGCGAGAGACCC	13111
QY	1261	GTGAAGCCCTGGCAGCTCTCTGAGAACTATGATCAACCTACCTCAGCTGGGCGGGCTG	13200
Db	1312	GTGAAGCCCTGGCAGCTCTCTGAGAACTATGATCAACCTACCTCAGCTGGGCGGGCTG	13711
QY	1321	CCGGCGGGCTTCGACACACACCGGAAACGTGGACATGATGAAGTAAAGCTGAAGTGTGGG	13800
Db	1372	CCGGCGGGCTTCGACACACACCGGAAACGTGGACATGATGAAGTAAAGCTGAAGTGTGGG	14311
QY	1381	TGGCAGGGCTCAGTGCCACAGGCTCCACAGAGTGGGAGGTTCAAGCGCAGCCTCAGAGCA	14400
Db	1432	TGGCAGGGCTCAGTGCCACAGGCTCCACAGAGTGGGAGGTTCAAGCGCAGCCTCAGAGCA	14911
QY	1441	GAGCGCCGTAAGATCCGCTGGACACGTCGTGAACAACCAAGACCGGTGACCCGGTGTGCG	15000
Db	1492	GAGCGCCGTAAGATCCGCTGGACACGTCGTGAACAACCAAGACCGGTGACCCGGTGTGCG	15511
QY	1501	CGGCGATGCCAGGAGGGGACAGGTGTCCGGGGGTTCAAGGGTTCCACTCCTGCTGTACACAC	15600
Db	1552	CGGCGATGCCAGGAGGGGACAGGTGTCCGGGGGTTCAAGGGTTCCACTCCTGCTGTACACAC	16111
QY	1561	TGTGTGACCTCGAGGGGGGACGCTACCGGCAAAACCCAGACGACATGCGCTGACCTTT	16200
Db	1612	TGTGTGACCTCGAGGGGGGACGCTACCGGCAAAACCCAGACGACATGCGCTGACCTTT	16711
QY	1621	TGTGGCCAGATGTGATGTGTCCCGCGAGCCGAAGACACGCTGCTCCGCGCAGAGGTCTGG	16800
Db	1672	TGTGGCCAGATGTGATGTGTCCCGCGAGCCGAAGACACGCTGCTCCGCGCAGAGGTCTGG	17311
QY	1681	TTTCTGGCATGGGGCGAGCCGGGCTGTGTGCTGCTCTCTCTGCTGTGAAGCTTGCGCTG	17400
Db	1732	TTTCTGGCATGGGGCGAGCCGGGCTGTGTGCTGCTCTCTCTGCTGTGAAGCTTGCGCTG	17911
QY	1741	GGCCTTGTGCTGGGTGTGTTGGGGGCTGTTCATACATCGGGAACGCCACTGTGTAG	18000
Db	1792	GGCCTTGTGCTGGGTGTGTTGGGGGCTGTTCATACATCGGGAACGCCACTGTGTAG	18511
QY	1801	GGCTCGGGGGGGCCCTCGGCGCTGCTTTTGGGCTGTGTGCTGGGCTGTGTGCTGCTCAGC	18600
Db	1852	GGCTCGGGGGGGGGCCCTCGGCGCTGCTTTTGGGCTGTGTGCTGGGCTGTGTGCTGCTCAGC	19111
QY	1861	GTCTCTCTGTCCCTGGGCCAGGCCAGGCCCTGCCCCGATGCTGTGACCAGCAGCCCTTTGTC	19200
Db	1912	GTCTCTCTGTCCCTGGGCCAGGCCAGGCCCTGCCCCGATGCTGTGACCAGCAGCCCTTTGTC	19711
QY	1921	CACGCCCGGCTCAAGGGGCTGCTGAGACACATTCCTCTCAAGGGCGCGAGATCTTGTGTG	19800
Db	1972	CACCTCCGCTCAGGGGCTGCTGTAGACACTCTTCTCTCAAGGGCGCGAGATCTTGTGTG	20311
QY	1981	GAGTCAGAACTGCTCTGAGCTGGGACAGACCGGCTGATGAGTGGCTGTGGGGGGGCTGTGG	20400

[illegible]

```

01      1 ATGCTGGGCGCTGCTGCTGGGCGCTGACCTGCGGCTCTCTGACACCTGGAGCGGG 60
02
03      99.9%: Score 255.8; DB 10; Length 3489;
04      Best Local Similarity 99.9%; Pred. No. 0;
05      Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
06
07      RESULT 9
08      US-09-796-338A-13
09      ; Sequence 13, Application US/09796338A
10      ; Patent No. US20020061522A1
11      ; GENERAL INFORMATION:
12      ; APPLICANT: Millennium Pharmaceuticals, Inc.
13      ; TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
14      ; TITLE OF INVENTION: 52872, NOVEL G PROTEIN-COUPLED RECEPTORS AND USES THEREFOR
15      ; FILE REFERENCE: 10448-020001
16      ; CURRENT APPLICATION NUMBER: US/09/796,338A
17      ; CURRENT FILING DATE: 2001-02-28
18      ; PRIOR APPLICATION NUMBER: US 60/186,059
19      ; PRIOR FILING DATE: 2000-02-29
20      ; NUMBER OF SEQ ID NOS: 26
21      ; SOFTWARE: FastSeq for Windows Version 4.0
22      ; SEQ ID NO 13
23      ; LENGTH: 3489
24      ; TYPE: DNA
25      ; ORGANISM: Homo sapiens
26      ; FEATURE:
27      ; NAME/KEY: CDS
28      ; LOCATION: (52)..(2607)
29      ; NAME/KEY: misc_feature
30      ; LOCATION: (1)..(3489)
31      ; OTHER INFORMATION: n = A,T,C or G
32      US-09-796-338A-13
33
34      Query Match 99.9%; Score 255.8; DB 10; Length 3489;
35      Best Local Similarity 99.9%; Pred. No. 0;
36      Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
37
38      1 ATGCTGGGCGCTGCTGCTGGGCGCTGACCTGCGGCTCTCTGACACCTGGAGCGGG 60
39
40      99.9%: Score 255.8; DB 10; Length 3489;
41      Best Local Similarity 99.9%; Pred. No. 0;
42      Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
43
44      RESULT 9
45      US-09-796-338A-13
46      ; Sequence 13, Application US/09796338A
47      ; Patent No. US20020061522A1
48      ; GENERAL INFORMATION:
49      ; APPLICANT: Millennium Pharmaceuticals, Inc.
50      ; TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
51      ; TITLE OF INVENTION: 52872, NOVEL G PROTEIN-COUPLED RECEPTORS AND USES THEREFOR
52      ; FILE REFERENCE: 10448-020001
53      ; CURRENT APPLICATION NUMBER: US/09/796,338A
54      ; CURRENT FILING DATE: 2001-02-28
55      ; PRIOR APPLICATION NUMBER: US 60/186,059
56      ; PRIOR FILING DATE: 2000-02-29
57      ; NUMBER OF SEQ ID NOS: 26
58      ; SOFTWARE: FastSeq for Windows Version 4.0
59      ; SEQ ID NO 13
60      ; LENGTH: 3489
61      ; TYPE: DNA
62      ; ORGANISM: Homo sapiens
63      ; FEATURE:
64      ; NAME/KEY: CDS
65      ; LOCATION: (52)..(2607)
66      ; NAME/KEY: misc_feature
67      ; LOCATION: (1)..(3489)
68      ; OTHER INFORMATION: n = A,T,C or G
69      US-09-796-338A-13
70
71      Query Match 99.9%; Score 255.8; DB 10; Length 3489;
72      Best Local Similarity 99.9%; Pred. No. 0;
73      Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
74
75      1 ATGCTGGGCGCTGCTGCTGGGCGCTGACCTGCGGCTCTCTGACACCTGGAGCGGG 60
76
77      99.9%: Score 255.8; DB 10; Length 3489;
78      Best Local Similarity 99.9%; Pred. No. 0;
79      Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
80
81      RESULT 9
82      US-09-796-338A-13
83      ; Sequence 13, Application US/09796338A
84      ; Patent No. US20020061522A1
85      ; GENERAL INFORMATION:
86      ; APPLICANT: Millennium Pharmaceuticals, Inc.
87      ; TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
88      ; TITLE OF INVENTION: 52872, NOVEL G PROTEIN-COUPLED RECEPTORS AND USES THEREFOR
89      ; FILE REFERENCE: 10448-020001
90      ; CURRENT APPLICATION NUMBER: US/09/796,338A
91      ; CURRENT FILING DATE: 2001-02-28
92      ; PRIOR APPLICATION NUMBER: US 60/186,059
93      ; PRIOR FILING DATE: 2000-02-29
94      ; NUMBER OF SEQ ID NOS: 26
95      ; SOFTWARE: FastSeq for Windows Version 4.0
96      ; SEQ ID NO 13
97      ; LENGTH: 3489
98      ; TYPE: DNA
99      ; ORGANISM: Homo sapiens
100     ; FEATURE:
101     ; NAME/KEY: CDS
102     ; LOCATION: (52)..(2607)
103     ; NAME/KEY: misc_feature
104     ; LOCATION: (1)..(3489)
105     ; OTHER INFORMATION: n = A,T,C or G
106     US-09-796-338A-13
107
108     Query Match 99.9%; Score 255.8; DB 10; Length 3489;
109     Best Local Similarity 99.9%; Pred. No. 0;
110     Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
111
112     1 ATGCTGGGCGCTGCTGCTGGGCGCTGACCTGCGGCTCTCTGACACCTGGAGCGGG 60
113
114     99.9%: Score 255.8; DB 10; Length 3489;
115     Best Local Similarity 99.9%; Pred. No. 0;
116     Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
117
118     RESULT 9
119     US-09-796-338A-13
120     ; Sequence 13, Application US/09796338A
121     ; Patent No. US20020061522A1
122     ; GENERAL INFORMATION:
123     ; APPLICANT: Millennium Pharmaceuticals, Inc.
124     ; TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
125     ; TITLE OF INVENTION: 52872, NOVEL G PROTEIN-COUPLED RECEPTORS AND USES THEREFOR
126     ; FILE REFERENCE: 10448-020001
127     ; CURRENT APPLICATION NUMBER: US/09/796,338A
128     ; CURRENT FILING DATE: 2001-02-28
129     ; PRIOR APPLICATION NUMBER: US 60/186,059
130     ; PRIOR FILING DATE: 2000-02-29
131     ; NUMBER OF SEQ ID NOS: 26
132     ; SOFTWARE: FastSeq for Windows Version 4.0
133     ; SEQ ID NO 13
134     ; LENGTH: 3489
135     ; TYPE: DNA
136     ; ORGANISM: Homo sapiens
137     ; FEATURE:
138     ; NAME/KEY: CDS
139     ; LOCATION: (52)..(2607)
140     ; NAME/KEY: misc_feature
141     ; LOCATION: (1)..(3489)
142     ; OTHER INFORMATION: n = A,T,C or G
143     US-09-796-338A-13
144
145     Query Match 99.9%; Score 255.8; DB 10; Length 3489;
146     Best Local Similarity 99.9%; Pred. No. 0;
147     Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
148
149     1 ATGCTGGGCGCTGCTGCTGGGCGCTGACCTGCGGCTCTCTGACACCTGGAGCGGG 60
150
151     99.9%: Score 255.8; DB 10; Length 3489;
152     Best Local Similarity 99.9%; Pred. No. 0;
153     Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
154
155     RESULT 9
156     US-09-796-338A-13
157     ; Sequence 13, Application US/09796338A
158     ; Patent No. US20020061522A1
159     ; GENERAL INFORMATION:
160     ; APPLICANT: Millennium Pharmaceuticals, Inc.
161     ; TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
162     ; TITLE OF INVENTION: 52872, NOVEL G PROTEIN-COUPLED RECEPTORS AND USES THEREFOR
163     ; FILE REFERENCE: 10448-020001
164     ; CURRENT APPLICATION NUMBER: US/09/796,338A
165     ; CURRENT FILING DATE: 2001-02-28
166     ; PRIOR APPLICATION NUMBER: US 60/186,059
167     ; PRIOR FILING DATE: 2000-02-29
168     ; NUMBER OF SEQ ID NOS: 26
169     ; SOFTWARE: FastSeq for Windows Version 4.0
170     ; SEQ ID NO 13
171     ; LENGTH: 3489
172     ; TYPE: DNA
173     ; ORGANISM: Homo sapiens
174     ; FEATURE:
175     ; NAME/KEY: CDS
176     ; LOCATION: (52)..(2607)
177     ; NAME/KEY: misc_feature
178     ; LOCATION: (1)..(3489)
179     ; OTHER INFORMATION: n = A,T,C or G
180     US-09-796-338A-13
181
182     Query Match 99.9%; Score 255.8; DB 10; Length 3489;
183     Best Local Similarity 99.9%; Pred. No. 0;
184     Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
185
186     1 ATGCTGGGCGCTGCTGCTGGGCGCTGACCTGCGGCTCTCTGACACCTGGAGCGGG 60
187
188     99.9%: Score 255.8; DB 10; Length 3489;
189     Best Local Similarity 99.9%; Pred. No. 0;
190     Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
191
192     RESULT 9
193     US-09-796-338A-13
194     ; Sequence 13, Application US/09796338A
195     ; Patent No. US20020061522A1
196     ; GENERAL INFORMATION:
197     ; APPLICANT: Millennium Pharmaceuticals, Inc.
198     ; TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
199     ; TITLE OF INVENTION: 52872, NOVEL G PROTEIN-COUPLED RECEPTORS AND USES THEREFOR
200     ; FILE REFERENCE: 10448-020001
201     ; CURRENT APPLICATION NUMBER: US/09/796,338A
202     ; CURRENT FILING DATE: 2001-02-28
203     ; PRIOR APPLICATION NUMBER: US 60/186,059
204     ; PRIOR FILING DATE: 2000-02-29
205     ; NUMBER OF SEQ ID NOS: 26
206     ; SOFTWARE: FastSeq for Windows Version 4.0
207     ; SEQ ID NO 13
208     ; LENGTH: 3489
209     ; TYPE: DNA
210     ; ORGANISM: Homo sapiens
211     ; FEATURE:
212     ; NAME/KEY: CDS
213     ; LOCATION: (52)..(2607)
214     ; NAME/KEY: misc_feature
215     ; LOCATION: (1)..(3489)
216     ; OTHER INFORMATION: n = A,T,C or G
217     US-09-796-338A-13
218
219     Query Match 99.9%; Score 255.8; DB 10; Length 34
```


Db 52 ATGCTGGGCCCTGCTGCTCGGGGCTCAGCCCTGCGGCTCTTCCTGACCCCTGGAGCGGG 111
Qy 61 GCCCATTTGTCGCTGCTCAGCAGCACTTAGATGAAGGGGAGACTAGCTGTGGGGGGCTG 120
Db 112 GCCCATTTGTCGCTGCTCAGCAGCACTTAGATGAAGGGGAGACTAGCTGTGGGGGGCTG 171
Qy 121 TTTCCCTGGGGGAGGGGAGGGGCTGCGCTCCGACCGGACAGGGCCAGCAGGCCCT 180
Db 172 TTTCCCTGGGGGAGGGGAGGGGCTGCGCTCCGACCGGACAGGGCCAGCAGGCCCT 231
Qy 181 GTGTGACACAGGTTCTCTCTCAAAACGGCTGCTGTGGGACATGGAATGGCCGTG 240
Db 222 GTGTGACACAGGTTCTCTCTCAAAACGGCTGCTGTGGGACATGGAATGGCCGTG 291
Qy 241 GAGGAGATCAACAAAGTCGGATGCTGCGCCGGGCTGCGGCTGGGCTAGCAGCTCTTT 300
Db 292 GAGGAGATCAACAAAGTCGGATGCTGCGCCGGGCTGCGGCTGGGCTAGCAGCTCTTT 351
Qy 301 GATACGTGCTGGAGCCGTGTGGTGGCCATGAAGCCAGCCATCTGCTGTGGCAAGCA 360
Db 352 GATACGTGCTGGAGCCGTGTGGTGGCCATGAAGCCAGCCATCTGCTGTGGCAAGCA 411
Qy 361 GGCAGCCGCGAGCATGCGCCCTACTGCAACTACAGCAGTACACAGCCCGCTGCTGGCT 420
Db 412 GGCAGCCGCGAGCATGCGCCCTACTGCAACTACAGCAGTACACAGCCCGCTGCTGGCT 471
Qy 421 GTCATCGGGGCCCACTGCTCAGAGCTGCGCATGTACCGGCAATTTCTTCAGTCTTTC 480
Db 472 GTCATCGGGGCCCACTGCTCAGAGCTGCGCATGTACCGGCAATTTCTTCAGTCTTTC 531
Qy 481 CTCATGCCCCAGTCAGTACGGTGTCTAGCATGAGCTGCTGAGCGCCCGGGAGACCTTC 540
Db 532 CTCATGCCCCAGTCAGTACGGTGTCTAGCATGAGCTGCTGAGCGCCCGGGAGACCTTC 591
Qy 541 CCTCTCTTCTCCGACCGCTGCGCCAGCGAGCTGTGACAGTACAGCGCCCGCGAGAGCTG 600
Db 592 CCTCTCTTCTCCGACCGCTGCGCCAGCGAGCTGTGACAGTACAGCGCCCGCGAGAGCTG 651
Qy 601 CTGCAAGAGTTGCGCTGGAACATGAGTGGCGCCCTGGGAGGAGCAGAGTACGGCCG 660
Db 652 CTGCAAGAGTTGCGCTGGAACATGAGTGGCGCCCTGGGAGGAGCAGAGTACGGCCG 711
Qy 661 CAGGAGCTGAGCATTTCTCGGCCCTGGCCGCGGACAGCGGCGATGTGATCGCGACAG 720
Db 712 CAGGAGCTGAGCATTTCTCGGCCCTGGCCGCGGACAGCGGCGATGTGATCGCGACAG 771
Qy 721 GGCCTGAGGCGCTGCGCCCGTGCAGTACCTGCGGGCTGGGGAAAGTGAAGACCTCTG 780
Db 772 GGCCTGAGGCGCTGCGCCCGTGCAGTACCTGCGGGCTGGGGAAAGTGAAGACCTCTG 831
Qy 781 CACCAAGTGAACAGAGCAGCTGACAGTGTGTGCTGTTCGCCCTGCGTGACCGCGCC 840
Db 832 CACCAAGTGAACAGAGCAGCTGACAGTGTGTGCTGTTCGCCCTGCGTGACCGCGCC 891
Qy 841 CACGCGCTTCTTCACTACATGATCAGACAGAGCTGTGCGCCAGAGTGTGGTGGCCAGC 900
Db 892 CACGCGCTTCTTCACTACATGATCAGACAGAGCTGTGCGCCAGAGTGTGGTGGCCAGC 951
Qy 901 GAGGAGCTGAGCACTGACCTGTGATGGGGGCTGGCGGAGTGGCCAGATGGGAGTGGGACG 960
Db 952 GAGGAGCTGAGCACTGACCTGTGATGGGGGCTGGCGGAGTGGCCAGATGGGAGTGGGACG 1011
Qy 961 GTGCTTGGCTTCTCAGAGGGGGTCCAGCTGACAGAGTTCGCCAGTACGTGAAGAG 1020
Db 1012 GTGCTTGGCTTCTCAGAGGGGGTCCAGCTGACAGAGTTCGCCAGTACGTGAAGAG 1071
Qy 1021 CACGTGGCCCTGGGCAACCGACCCGGCTTCTGCTGCTGGCGGAGAGAGGAGAGGCT 1080
Db 1072 CACGTGGCCCTGGGCAACCGACCCGGCTTCTGCTGCTGGCGGAGAGAGGAGAGGCT 1131
Qy 1081 CTGGAGGAGAGAGCTGTGGGGGAGGCTGCGGAGAGTGTGATGATCAAGCTGCAAGAC 1140
Db 1132 CTGGAGGAGAGAGCTGTGGGGGAGGCTGCGGAGAGTGTGATGATCAAGCTGCAAGAC 1191

Qy 1141 GTGAGCGAGGGGCTAAATCACCAGACGTTCTGTGTACAGCAGTGTATAGCGG 1200
Db 1192 GTGAGCGAGGGGCTAAATCACCAGACGTTCTGTGTACAGCAGTGTATAGCGG 1251
Qy 1201 GCCAGGGCCCTGACACAACTTTCAGTGCACAGCCCTCAGGCTGCCCGGAGAGACCC 1260
Db 1252 GCCAGGGCCCTGACACAACTTTCAGTGCACAGCCCTCAGGCTGCCCGGAGAGACCC 1311
Qy 1261 GTGAAGCCCTGGCAGCTCTGTGAGAACTGTAAACCTGACTTTCACAGTGGGGGGCTG 1320
Db 1312 GTGAAGCCCTGGCAGCTCTGTGAGAACTGTAAACCTGACTTTCACAGTGGGGGGCTG 1371
Qy 1321 CCGCTGCGGTTGAGACAGCAGCGGAAACGTGACATGAGATACAGCTGAGCTGTGGGT 1380
Db 1372 CCGCTGCGGTTGAGACAGCAGCGGAAACGTGAGATGAGATACAGCTGAGCTGTGGGT 1431
Qy 1381 TGGCAGGGCTCAGTGCCTCCAGGCTCCAGAGCTGGGAGGTTCAAGCGCAGCTCAGAGCA 1440
Db 1432 TGGCAGGGCTCAGTGCCTCCAGGCTCCAGAGCTGGGAGGTTCAAGCGCAGCTCAGAGCA 1491
Qy 1441 GAGCGCTGGAATCCGCTGGGAGACGCTGTGACAAACCAAGAACCCGCTGTCGGGTCTG 1500
Db 1492 GAGCGCTGGAATCCGCTGGGAGACGCTGTGACAAACCAAGAACCCGCTGTCGGGTCTG 1551
Qy 1501 CCGCAGTGCAGAGAGGGCCAGGCTGCGCGGGGTCAAGAGGGTTCCACTCTGCTTACAGAC 1560
Db 1552 CCGCAGTGCAGAGAGGGCCAGGCTGCGCGGGGTCAAGAGGGTTCCACTCTGCTTACAGAC 1611
Qy 1561 TGTGTGACTGCGAGGCGGGGACGCTACCGGCAAAACCCAGAGACATCGCTGACCTTT 1620
Db 1612 TGTGTGACTGCGAGGCGGGGACGCTACCGGCAAAACCCAGAGACATCGCTGACCTTT 1671
Qy 1621 TGTGGCCAGATGAGTGTGTCCCGGAGGAGAACACACAGTGTGTCCCGCGAGGCTGCG 1680
Db 1672 TGTGGCCAGATGAGTGTGTCCCGGAGGAGAACACACAGTGTGTCCCGCGAGGCTGCG 1731
Qy 1681 TTTCTGAGGAGGGGAGAGCGGCTGTGCTGCTGCTGCTCTCTGAGAGCTGAGCGCTG 1740
Db 1732 TTTCTGAGGAGGGGAGAGCGGCTGTGCTGCTGCTGCTCTCTGAGAGCTGAGCGCTG 1791
Qy 1741 GGCCTTGTGTGCTGCTTGTGGGGCTGTTTCACACATCGGAGACGCCACTGTTCA 1800
Db 1792 GGCCTTGTGTGCTGCTTGTGGGGCTGTTTCACACATCGGAGACGCCACTGTTCA 1851
Qy 1801 GGCCTGAGGGGGGCGCCCTGGCCGCTTGTGGCGTGTGTGCTGCGGCTGCTGCGCTAC 1860
Db 1852 GGCCTGAGGGGGGCGCCCTGGCCGCTTGTGGCGTGTGTGCTGCGGCTGCTGCGCTAC 1911
Qy 1861 GTCTCTGTGTTCCCTGTGGCAGCCAGCCCTGCGGATGCTGCGGACAGCCCTTGTCC 1920
Db 1912 GTCTCTGTGTTCCCTGTGGCAGCCAGCCCTGCGGATGCTGCGGACAGCCCTTGTCC 1971
Qy 1921 CACCTCCCGCTACAGGGCTGCTGAGACACACTCTTCTGCAAGCGGCGGAGATCTTCTG 1980
Db 1972 CACCTCCCGCTACAGGGCTGCTGAGACACACTCTTCTGCAAGCGGCGGAGATCTTCTG 2031
Qy 1981 GAGTCAGAACTGCTCTGTGAGTGGGAGACGCGGCTGAGTGGGCTGCGTGGGGGGCCG 2040
Db 2032 GAGTCAGAACTGCTCTGTGAGTGGGAGACGCGGCTGAGTGGGCTGCGTGGGGGGCCG 2091
Qy 2041 GCTTGGCTGTGTGCTGCTGTGCGCATGTGTTGAGAGTGTGCACTGTGCACTGTGATCTG 2100
Db 2092 GCTTGGCTGTGTGCTGCTGTGCGCATGTGTTGAGAGTGTGCACTGTGCACTGTGATCTG 2151
Qy 2101 GTGGCCCTTCCCGCGAGGTGTGACAGTGTGACACATGCTGCCACGAGGCGCTGCTG 2160
Db 2152 GTGGCCCTTCCCGCGAGGTGTGACAGTGTGACAGTGTGACATGCTGCCACGAGGCGCTGCTG 2211
Qy 2161 CACTGCGGACAGAGCTTCCGGGTCAAGCTTCCGCTTACAGCAGAGCCCAATGCTCAAGCTG 2220
Db 2212 CACTGCGGACAGAGCTTCCGGGTCAAGCTTCCGCTTACAGCAGAGCCCAATGCTCAAGCTG 2271

QY 2221 GCGTTTCCTGCTTCCTGCGGACATTTCTGCTGCGAGCCAGCCGGGCTGCTACACCGT 2280
 Db 2272 GCGTTTCCTGCTTCCTGCGGACATTTCTGCTGCGAGCCAGCCGGGCTGCTACACCGT 2331
 QY 2281 GCGGTTGGCTCAGCTTTGCGATCTGCGCTACTTCAATCACTGGGTCTCTTTGTGCCC 2340
 Db 2332 GCGGTTGGCTCAGCTTTGCGATCTGCGCTACTTCAATCACTGGGTCTCTTTGTGCCC 2391
 QY 2341 CTCTGCGCAATGTCAGGTGCTCCAGGCGCCGCTGCGATGGGCGCCCTCTCTCTC 2400
 Db 2392 CTCTGCGCAATGTCAGGTGCTCCAGGCGCCGCTGCGATGGGCGCCCTCTCTCTC 2451
 QY 2401 TGTGTCCTGGGCAATCTGCTGCTCCCTCCACCTGCGCAGGTGTACTCTCATGCGGAG 2460
 Db 2452 TGTGTCCTGGGCAATCTGCTGCTCCCTCCACCTGCGCAGGTGTACTCTCATGCGGAG 2511
 QY 2461 CCAAGGCTCAACACCCCGAGTTCTCTGCGAGGGGGCCCTGGGAGTGGCCAGAGCCAG 2520
 Db 2512 CCAAGGCTCAACACCCCGAGTTCTCTGCGAGGGGGCCCTGGGAGTGGCCAGAGCCAG 2571
 QY 2521 AATGACGGGAACACAGGAATCAAGGGGAACATGATGA 2559
 Db 2572 AATGACGGGAACACAGGAATCAAGGGGAACATGATGA 2610

RESULT 10

US-09-799-629-3
 ; Sequence 3, Application US/09799629
 ; Publication No. US2003008344A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ADLER, JON ELLIOT
 ; APPLICANT: ZOZULYA, SERGEY
 ; APPLICANT: LI, XIADONG
 ; APPLICANT: O'CONNELL, SHAWN
 ; APPLICANT: STASZEMSKI, LENA
 ; TITLE OF INVENTION: TLR TASTE RECEPTORS AND GENES ENCODING SAME
 ; FILE REFERENCE: 078003/0277870/RXT
 ; CURRENT APPLICATION NUMBER: US/09/799,629
 ; PRIOR FILING DATE: 2001-03-07
 ; PRIOR APPLICATION NUMBER: 60/187,546
 ; PRIOR FILING DATE: 2000-03-07
 ; PRIOR APPLICATION NUMBER: 60/195,536
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: 60/209,840
 ; PRIOR FILING DATE: 2000-06-06
 ; PRIOR APPLICATION NUMBER: 60/214,213
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: 60/226,448
 ; PRIOR FILING DATE: 2000-08-17
 ; PRIOR APPLICATION NUMBER: 60/259,227
 ; PRIOR FILING DATE: 2001-01-03
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 2553
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-799-629-3

Query Match 99.1%; Score 2537; DB 9; Length 2553;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2553; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY 1 ATGCTGGGCGCTGCTGCTGCGGCTCAGGCTTGCGCTCTCTGCAACCTGAGGAGGG 60
 Db 1 ATGCTGGGCGCTGCTGCTGCGGCTCAGGCTTGCGCTCTCTGCAACCTGAGGAGGG 60
 QY 61 GCCCATTTGCTGTGTCACAGCACTTAAAGATGAAGGGGAGTAAGTCTGCTGGGGGCTG 120
 Db 61 GCCCATTTGCTGTGTCACAGCACTTAAAGATGAAGGGGAGTAAGTCTGCTGGGGGCTG 120
 QY 121 TTCCCTTGGGCGAGGCGGAGAGGCTGCGCTCCGACCGCGGACAGGCCAGAGCCCT 180
 Db 121 TTCCCTTGGGCGAGGCGGAGAGGCTGCGCTCCGACCGCGGACAGGCCAGAGCCCT 180

Db 121 TTCCCTTGGGCGAGGCGGAGAGGCTGCGCTCCGACCGCGGACAGGCCAGAGCCCT 180
 QY 181 GTGAGCAACAGGTTCTCTCAAAAGGCGCTGCTGGGCACTGGGCAATGAATAGGCGGT 240
 Db 181 GTGAGCAACAGGTTCTCTCAAAAGGCGCTGCTGGGCACTGGGCAATGAATAGGCGGT 240
 QY 241 GAGAGATCAACAACAAGTGGATCTGCTCCGCGGTGGCGCTGGGCTACAGACTCTTT 300
 Db 241 GAGAGATCAACAACAAGTGGATCTGCTCCGCGGTGGCGCTGGGCTACAGACTCTTT 300
 QY 301 GATACGTCTGCGAGCTGTGTGCGCATATGAGGCCACAGCTTATCTCTGCGGCAAGGCA 360
 Db 301 GATACGTCTGCGAGCTGTGTGCGCATATGAGGCCACAGCTTATCTCTGCGGCAAGGCA 360
 QY 361 GGCAGCGCGCAATCGCGGCTCTACTGCAATACAGAGTACAGAGCCCGGTGTGCTGCT 420
 Db 361 GGCAGCGCGCAATCGCGGCTCTACTGCAATACAGAGTACAGAGCCCGGTGTGCTGCT 420
 QY 421 GTCATCGGGCCCACTGCTGAGAGCTTGCATAGGTCACCGGCAAGTCTTCAAGCTTCTC 480
 Db 421 GTCATCGGGCCCACTGCTGAGAGCTTGCATAGGTCACCGGCAAGTCTTCAAGCTTCTC 480
 QY 481 CTATGCGCCCAAGTCACTACGCTGATGAGTGTGATGAGTGTGAGCGCCGCGGAGACTTC 540
 Db 481 CTATGCGCCCAAGTCACTACGCTGATGAGTGTGATGAGTGTGAGCGCCGCGGAGACTTC 540
 QY 541 CCTCTCTTCTTCCGACCGTGCAGCGGACCGTGTGAGCTGACGCGCGCGCGGAGACTG 600
 Db 541 CCTCTCTTCTTCCGACCGTGCAGCGGACCGTGTGAGCTGACGCGCGCGCGGAGACTG 600
 QY 595 CTGACAGAGTTCGGCTGGAACCTGGGTGGCGGCTGCGGACGCGACGAGTACGGCGG 660
 Db 595 CTGACAGAGTTCGGCTGGAACCTGGGTGGCGGCTGCGGACGCGACGAGTACGGCGG 660
 QY 661 CAGGCGCTGAGCAATCTCTGCGGCTGCGCGGCGGACGCGGCAATCTGATCGGCGAGAG 720
 Db 661 CAGGCGCTGAGCAATCTCTGCGGCTGCGCGGCGGACGCGGCAATCTGATCGGCGAGAG 720
 QY 721 GCGCTGTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
 Db 721 GCGCTGTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 780
 QY 781 CACCAAGTGAACACAGAGACAGCTGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 781 CACCAAGTGAACACAGAGACAGCTGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 QY 841 CAGGCGCTTCAACATCACTGATCAGACAGAGGCTGCGGCGGCGGCGGCGGCGGCGGCGG 900
 Db 841 CAGGCGCTTCAACATCACTGATCAGACAGAGGCTGCGGCGGCGGCGGCGGCGGCGGCGG 900
 QY 895 GAGGCGTGGCTGACCTGTGACCTGTGATGAGGCGTGGCGGCGGCGGCGGCGGCGGCGG 954
 Db 895 GAGGCGTGGCTGACCTGTGACCTGTGATGAGGCGTGGCGGCGGCGGCGGCGGCGGCGG 954
 QY 961 GTGCTTGGCTTCTCTCCAGAGAGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
 Db 961 GTGCTTGGCTTCTCTCCAGAGAGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
 QY 1021 CACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 Db 1021 CACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 QY 1081 CTGAGAGAGAGAGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
 Db 1081 CTGAGAGAGAGAGTGTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
 QY 1141 GTGAGCGCAGGCGCTAAATCAACACAGAGCTTCTGCTTACAGAGCTGTGTATAGCTG 1200
 Db 1141 GTGAGCGCAGGCGCTAAATCAACACAGAGCTTCTGCTTACAGAGCTGTGTATAGCTG 1200
 QY 1201 GCCCAGGCGCTGCAACAACATCTTCACTGATGAGGCGGCGGCGGCGGCGGCGGCGG 1260
 Db 1201 GCCCAGGCGCTGCAACAACATCTTCACTGATGAGGCGGCGGCGGCGGCGGCGGCGG 1260
 QY 1195 GCCCAGGCGCTGCAACAACATCTTCACTGATGAGGCGGCGGCGGCGGCGGCGGCGG 1254
 Db 1195 GCCCAGGCGCTGCAACAACATCTTCACTGATGAGGCGGCGGCGGCGGCGGCGGCGG 1254

QY 1261 GTGAAGCCCTGGACAGCTCTTGAGAAACATGTACACCTGACCTTCCACGTGGCGGGCTG 1320
 |||||
 Db 1255 GTGAAGCCCTGGACAGCTCTTGAGAAACATGTACACCTGACCTTCCACGTGGCGGGCTG 1314
 QY 1321 CCGCTGGCGGTTGACAGACAGCGGAAACGTGACATGAGTACGACCTGAAAGCTGTGGGTG 1380
 |||||
 Db 1315 CCGCTGGCGGTTGACAGACAGCGGAAACGTGACATGAGTACGACCTGAAAGCTGTGGGTG 1374
 QY 1381 TGGCAGGGGCTCAGTCCGACAGGCTCCACAGACGTTGAGTTCACAGGCAAGCTCTAGAGCA 1440
 |||||
 Db 1375 TGGCAGGGGCTCAGTCCGACAGGCTCCACAGACGTTGAGTTCACAGGCAAGCTCTAGAGCA 1434
 QY 1441 GAGCGCCCTGAAGATCCGCTGGACACGCTCTGACAAACCAAGACCCGCTGCCGAGTCTG 1500
 |||||
 Db 1435 GAGCGCCCTGAAGATCCGCTGGACACGCTCTGACAAACCAAGACCCGCTGCCGAGTCTG 1494
 QY 1501 CGGCACTGACAGAGAGGCGCAGGTCGCGGGGTCAAGGGGGTTCACCTCGCTGTACGAC 1560
 |||||
 Db 1495 CGGCACTGACAGAGAGGCGCAGGTCGCGGGGTCAAGGGGGTTCACCTCGCTGTACGAC 1554
 QY 1561 TGTGTGAGCTGGAGGCGGAGGCTACGCGGCAAAACCCAGACGATGGCCCTGACCTTT 1620
 |||||
 Db 1555 TGTGTGAGCTGGAGGCGGAGGCTACGCGGCAAAACCCAGACGATGGCCCTGACCTTT 1614
 QY 1621 TGTGCGCAGATGAGTGGTCCCGGAGCGAAGACACACGCTGTCGCGCGAGGTCTCG 1680
 |||||
 Db 1615 TGTGCGCAGATGAGTGGTCCCGGAGCGAAGACACACGCTGTCGCGCGAGGTCTCG 1674
 QY 1681 TTCTGCGCATGGGCGGAGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
 |||||
 Db 1675 TTCTGCGCATGGGCGGAGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1734
 QY 1741 GAGCCCTTGCTGGGCTTTTGGGGGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
 |||||
 Db 1735 GAGCCCTTGCTGGGCTTTTGGGGGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1794
 QY 1801 GCGTGG 1860
 |||||
 Db 1795 GCGTGG 1854
 QY 1861 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
 |||||
 Db 1855 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1914
 QY 1921 CACCTCCGCTACAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
 |||||
 Db 1915 CACCTCCGCTACAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1974
 QY 1981 GAGTCAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
 |||||
 Db 1975 GAGTCAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2034
 QY 2041 GCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
 |||||
 Db 2035 GCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2094
 QY 2101 GTGGGCTTCCGCGGAGGAGTGTGAGAGACGTCGACATGCTCCACAGGAGGCGTGGGTG 2160
 |||||
 Db 2095 GTGGGCTTCCGCGGAGGAGTGTGAGAGACGTCGACATGCTCCACAGGAGGCGTGGGTG 2154
 QY 2161 CACTGCGCAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
 |||||
 Db 2155 CACTGCGCAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2214
 QY 2221 GCGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
 |||||
 Db 2215 GCGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2274
 QY 2281 GCGGCTGAGCCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
 |||||
 Db 2275 GCGGCTGAGCCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2334

QY 2341 CTCTGGCCAAATGTGACAGTGTGCTCTCAGGCCCGCCGCTGACAGATGGCGCCCTCTGCTG 2400
 |||||
 Db 2335 CTCTGGCCAAATGTGACAGTGTGCTCTCAGGCCCGCCGCTGACAGATGGCGCCCTCTGCTG 2394
 QY 2401 TGTGTCTGGGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
 |||||
 Db 2395 TGTGTCTGGGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2454
 QY 2461 CCAGGGCTCAACACCCCGAGTCTTCTGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2520
 |||||
 Db 2455 CCAGGGCTCAACACCCCGAGTCTTCTGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2514
 QY 2521 AATGACGGGAAACACAGAAATCAGGGGAAACATGAGTGA 2559
 |||||
 Db 2515 AATGACGGGAAACACAGAAATCAGGGGAAACATGAGTGA 2553
 RESULT 11
 US-09-927-315-13
 : Sequence 13, Application US/09927315
 : Publication No. US20030040045A1
 : GENERAL INFORMATION:
 : APPLICANT: Zuker, Charles S.
 : APPLICANT: Ryda, Nicholas J.P.
 : APPLICANT: Nelson, Greg
 : APPLICANT: Hoon, Mark A.
 : APPLICANT: Chandrasekar, Jayaram
 : APPLICANT: Zhang, Yifeng
 : APPLICANT: The Regents of the University of California
 : APPLICANT: The Government of the United States of America
 : APPLICANT: as represented by the Secretary of the
 : TITLE OF INVENTION: Mammalian Sweet Taste Receptors
 : FILE REFERENCE: 02307E-120110US
 : CURRENT APPLICATION NUMBER: US/09/927,315
 : PRIOR FILING DATE: 2001-08-10
 : PRIOR APPLICATION NUMBER: US 60/302,898
 : NUMBER OF SEQ ID NOS: 25
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO: 13
 : LENGTH: 3200
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : OTHER INFORMATION: human T1R3 sweet taste receptor genomic
 US-09-927-315-13
 Query Match 76.2%; Score 1950; DB 9; Length 3200;
 Best Local Similarity 82.1%; Pred. No. 0;
 Matches 2559; Conservative 0; Mismatches 0; Indels 559; Gaps 5;
 QY 1 ATGCTGGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
 |||||
 Db 67 ATGCTGGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 126
 QY 61 GCCCATTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 |||||
 Db 127 GCCCATTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 186
 QY 121 TTCCCGCTGGGCGAGGCGGAGGAGGCTGCTGCCAGCGGAGACAGGCGCCAGACAGCCCT 180
 |||||
 Db 187 TTCCCGCTGGGCGAGGCGGAGGAGGCTGCTGCCAGCGGAGACAGGCGCCAGAGCCCT 246
 QY 181 GTGTGCA----- 187
 |||||
 Db 247 GTGTGCAACAGGTACAGAGGTGGACGGCTGGGTGCTGGGGTCAAGGGTACAGGTCTGGG 306
 QY 188 -----CCAGGTTCTCT 199
 |||||
 Db 307 GTGTCTGAGCTGGGCGCGAGAGTGGCACTTGGCGTTCTGTGTGGCCCGCAGGTTCTCT 366
 |||||
 QY 200 CAACGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 259

Db 367 CAAACGGCGCTCTGGGCACTGGCCATGAATAATGGCCGTGGAGAGATCAACAACAAGT 426
 QY 260 CGATATGCTCCCGGGGTGGCCCTGAGACCTCTTTGATACGGCTCGGAGCCTG 319
 Db 427 CGGATCTGCTCCCGGGGTGGCCCTGAGACCTCTTTGATACGGCTCGGAGCCTG 486
 QY 320 TGTGGCCATGAAGCCAGCCTCATGTTCTCGGCCAAGGCGAGCGCCGACATCCGCG 379
 Db 487 TGTGGCCATGAAGCCAGCCTCATGTTCTCGGCCAAGGCGAGCGCCGACATCCGCG 546
 QY 380 CCTATGCACTACACGAGTACCAAGCCGCTGTGCTGTGATGAGGCCCCACTGCT 439
 Db 547 CCTATGCACTACACGAGTACCAAGCCGCTGTGCTGTGATGAGGCCCCACTGCT 606
 QY 440 CAGAGCTGCGCATGTCACCGGCAAGTTCTTCACTCTCTTCATATGCC----- 489
 Db 607 CAGAGCTGCGCATGTCACCGGCAAGTTCTTCACTCTCTTCATATGCC----- 666
 QY 490 ----- 489
 Db 667 CCCCCACATGACCCACCCACCCAGCCCTGCCGTGGAGCCCTGTGTACAGAGATG 726
 QY 490 -----CAGTCAAGTACAGGTGCTACATGAGCTGTGAGCCCGGGAGA 535
 Db 727 CCTCTTGGCCCTTGCAGGTACAGTACGTGTACATGAGCTGTGAGCCCGGGAGA 786
 QY 536 CCTTCCCTCTCTTCCGACCCGTCGACGACCCGTCGACGTGACGTGACGCGCCGCG 595
 Db 787 CCTTCCCTCTCTTCCGACCCGTCGACGACCCGTCGACGTGACGCGCCGCGCG 846
 QY 596 AGCTGTGCAAGAGATTGGGTGGAATGAGTGGCCCGCCCGGAGCAGCAGCAGTACG 655
 Db 847 AGCTGTGCAAGAGATTGGGTGGAATGAGTGGCCCGCCCGGAGCAGCAGCAGTACG 906
 QY 656 GCGGCGAGGGCTGAGATCTTCTCGGCCCTGGCCCGGGGACAGGGGATCTGCACTCGGC 715
 Db 907 GCGGCGAGGGCTGAGATCTTCTCGGCCCTGGCCCGGGGACAGGGGATCTGCACTCGGC 966
 QY 716 ACGAGGGCTGTGCTGCGCTGCGCCGTCGACGACGCTGCGGAGGAGTGAAGAGAG 775
 Db 967 ACGAGGGCTGTGCTGCGCTGCGCCGTCGACGACGCTGCGGAGGAGTGAAGAGAG 1026
 QY 776 TCTGTCAACAGGTGAACAGAGAGCGTGAAGTGTGCTGCTTCCGCTCCGTGACG 835
 Db 1027 TCTGTCAACAGGTGAACAGAGAGCGTGAAGTGTGCTGCTTCCGCTCCGTGACG 1086
 QY 836 CCGGCCAGGCCCTTCTCACTACAGCATCAGACAGGCTTCCGCCAAGGTGTGGGTGG 895
 Db 1087 CCGGCCAGGCCCTTCTCACTACAGCATCAGACAGGCTTCCGCCAAGGTGTGGGTGG 1146
 QY 896 CCAAGCGAGGCTGTGCTACCTGTGACGTGATGAGGAGCTCCGCGCATGGCCAGATG 955
 Db 1147 CCAAGCGAGGCTGTGCTACCTGTGACGTGATGAGGAGCTCCGCGCATGGCCAGATG 1206
 QY 956 GCAAGGCTGTGGCTTCTCTCAAGAGGGGTGCCAGCTGACAGAGTTCGCCAGTACGTGA 1015
 Db 1207 GCAAGGCTGTGGCTTCTCTCAAGAGGGGTGCCAGCTGACAGAGTTCGCCAGTACGTGA 1266
 QY 1016 AGAGCGACCTGTGGCCCTGACCCAGACCGGCGCTTGTGTCGCTGGGCGAGAGGGAGC 1075
 Db 1267 AGAGCGACCTGTGGCCCTGACCCAGACCGGCGCTTGTGTCGCTGGGCGAGAGGGAGC 1326
 QY 1076 AGGCTGTGAGAGAGAGCTGTGGGCCAGCGCTCCCGCATGTGATGATCACTACGCTGC 1135
 Db 1327 AGGCTGTGAGAGAGAGCTGTGGGCCAGCGCTCCCGCATGTGATGATCACTACGCTGC 1386
 QY 1136 AGAAGCTGAGGCGAGGCTTAATATACCAACAGAGCTTCTGTGTACGACGCTGTATA 1195
 Db 1387 AGAAGCTGAGGCGAGGCTTAATATACCAACAGAGCTTCTGTGTACGACGCTGTATA 1446
 QY 1196 GCGTGGGCGAGGCGCTGTGACAAACACTCTTCAAGTGAACGCTTACGCTGCCCGCGCAG 1255
 Db 1255 GCGTGGGCGAGGCGCTGTGACAAACACTCTTCAAGTGAACGCTTACGCTGCCCGCGCAG 1315

Db 1447 GCGTGGGCGAGGCGCTGTGACAAACACTCTTCAAGTGAACGCTTACGCTGCCCGCGCAG 1506
 QY 1256 ACCCGTGAAGCCCTG----- 1271
 Db 1507 ACCCGTGAAGCCCTGTGACAGGTGAGCCCGGAGATGGGGGTGTGCTGTCTCATATGTG 1566
 QY 1272 ----- 1271
 Db 1567 CCCAGGCCACAGGCGAGGCCACACGCTGAGCTGGAAGGTGGCTGGCGGCTGACGCCCG 1626
 QY 1272 -----GCACCTCTGGAAGAACATGTAAACCTGACCTTCCAGTGGGGGCTGCG 1321
 Db 1627 TCCCGCCCGCCGACCTCTGTGAAGAACATGTAAACCTGACCTTCCAGTGGGGGCTGCG 1686
 QY 1322 CGCTGCGGTTGACAGAGAGCGGAACGTCGACATGAGTACGACCTGAGCTGAGGTGGGTGT 1381
 Db 1687 CGCTGCGGTTGACAGAGAGCGGAACGTCGACATGAGTACGACCTGAGCTGAGGTGGGTGT 1746
 QY 1382 GGCAGGGCTCAGTGCACAGGCTCCACAGAGCTGGGCAAGTTCCAAAGCCCTCAGACAG 1441
 Db 1747 GGCAGGGCTCAGTGCACAGGCTCCACAGAGCTGGGCAAGTTCCAAAGCCCTCAGACAG 1806
 QY 1442 AGCGCTGAAGATCCGCTGGTGCACAGCTGTGACAC----- 1476
 Db 1807 AGCGCTGAAGATCCGCTGGTGCACAGCTGTGACACAGCTGAGGTGAGGTGGGTGTGCC 1866
 QY 1477 ----- 1476
 Db 1867 AGGCTGCCCGCTGTACGCCCGCGGCGAGCGGCGAGCCTGGGGGTGGGCGCTTCCAGT 1926
 QY 1477 -----CAGAGCCCGT 1487
 Db 1927 CTCCCGTGGGATCCAGCCAGACAGAGCCAGACCCAGCCTGTGCGAGAGGCCGT 1986
 QY 1488 GTCCCGGTGCTCGGGAGTCCAGAGAGGGGCGAGGTGCGCGGGTCCAGAGGGGCTTCCAGT 1547
 Db 1987 GTCCCGGTGCTCGGGAGTCCAGAGAGGGGCGAGGTGCGCGGGTCCAGAGGGGCTTCCAGT 2046
 QY 1548 CTGCTGTACAGCTGTGTGACCTGAGAGGGGCGAGGTACCGGCGCAAAAC----- 1597
 Db 2047 CTGCTGTACAGCTGTGTGACCTGAGAGGGGCGAGGTACCGGCGCAAAACAGGTGAGCC 2106
 QY 1598 ----- 1597
 Db 2107 GCTTCCCGGAGCGGGGCTGGGAGACAGACAGGAGGCTCTGCCAAGTCTGTGACTC 2166
 QY 1598 -----CAGA 1601
 Db 2167 TGAGACAGAGCCACAGGCTACAGAGCAACACCCAGCGCTTCTCTCTCTACACAGA 2226
 QY 1602 CGACATGCGCTGACCTTTTGTGGCCAGAGTGAATGTTCCCGGAGCGAAGACACAGCCTG 1661
 Db 2227 CGACATGCGCTGACCTTTTGTGGCCAGAGTGAATGTTCCCGGAGCGAAGACACAGCCTG 2286
 QY 1662 CTTCCGCGCAGAGTGTGCTGCTGAGATGGGGGAGCCGCTGTGCTGCTGCTCT 1721
 Db 2287 CTTCCGCGCAGAGTGTGCTGCTGAGATGGGGGAGCCGCTGTGCTGCTGCTCT 2346
 QY 1722 GCTGCTAGGCTGTGCGGCTGTTGTGCTGCTGCTGTTGGGGGCTGTTGCTACATG 1781
 Db 2347 GCTGCTAGGCTGTGCGGCTGTTGTGCTGCTGCTGTTGGGGGCTGTTGCTACATG 2406
 QY 1782 GGACAGCCCACTGTTAGGCTTGGGGGGGCGCCCTGCTGCTTGTGGCTGTGTGCT 1841
 Db 2407 GGACAGCCCACTGTTAGGCTTGGGGGGGCGCCCTGCTGCTTGTGGCTGTGTGCT 2466
 QY 1842 GGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1901
 Db 2467 GGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2526
 QY 1902 GGGCCAGACGCGCTTGTGCGACAGCTCCGCGTACAGGGGTGCTGAGCAGCTTCCGCA 1961
 Db 2527 GGGCCAGACGCGCTTGTGCGACAGCTCCGCGTACAGGGGTGCTGAGCAGCTTCCGCA 2586

OY 1962 GCGGCGCGAGATCTTGTGAGATCAGAACTGCTCTGAGCTGGGAGACCGGCTGAGTGG 2021
|||
DB 2587 GCGGCGCGAGATCTTGTGAGATCAGAACTGCTCTGAGCTGGGAGACCGGCTGAGTGG 2646
OY 2022 CTGCTGCGGGGCGGCTGGGCGTGGTGGTGTCTGTGACCATGCTGTGGAGATGCGC 2081
|||
DB 2647 CTGCTGCGGGGCGGCTGGGCGTGGTGGTGTCTGTGACCATGCTGTGGAGATGCGC 2706
OY 2082 ACTGTGACACCTGTACCTGTGGCTTCCCGCGGAGGTGTGAGAGGATGGCACATGCT 2141
|||
DB 2707 ACTGTGACACCTGTGTACCTGTGGCTTCCCGCGGAGGTGTGAGAGGATGGCACATGCT 2766
OY 2142 GCGGCGAGGAGGCTGTGTGACATGCGGACAGCTCTGTGGTGTGAGCTGTGGCTGAGCGCA 2201
|||
DB 2767 GCGGCGAGGAGGCTGTGTGACATGCGGACAGCTCTGTGGTGTGAGCTGTGGCTGAGCGCA 2826
OY 2202 GCGGCGCAATGCGGAGCTGTGGCTTCTGTCTGTCTGTGGCACTTCTGTGGTGGAGCGCA 2261
|||
DB 2827 GCGGCGCAATGCGGAGCTGTGGCTTCTGTCTGTCTGTGGCACTTCTGTGGTGGAGCGCA 2886
OY 2262 GCGGCGGCTGTGTACACCTGTGGCTGTGACCTTGTGACATGCTGTGGCTTGTATCAG 2321
|||
DB 2887 GCGGCGGCTGTGTACACCTGTGGCTGTGACCTTGTGACATGCTGTGGCTTGTATCAG 2946
OY 2322 CTGGGCTCTCTTGTGGCCCTCTGTGGCAATGTGAGGTGTCTGACGCGCGCGGTGCA 2381
|||
DB 2947 CTGGGCTCTCTTGTGGCCCTCTGTGGCAATGTGAGGTGTCTGACGCGCGCGGTGCA 3006
OY 2382 GATGGGCGCCCTCTGTGTGTGTCTGTGGGCAATGTGCTGTGGCTTGTCCACCTGCCAGGTG 2441
|||
DB 3007 GATGGGCGCCCTCTGTGTGTGTCTGTGGGCAATGTGCTGTGGCTTGTCCACCTGCCAGGTG 3066
OY 2442 TTACCTGTCATGCGGAGCGGAGCTCAACACCGCCGAGTCTTCTGTGGGAGGCGGCGCC 2501
|||
DB 3067 TTACCTGTCATGCGGAGCGGAGCTCAACACCGCCGAGTCTTCTGTGGGAGGCGGCGCC 3126
OY 2502 TGGGATGCGCCAAAGCGCAAGATGACGGGAAACAGAGAAATCATGAGATGTA 2559
|||
DB 3127 TGGGATGCGCCAAAGCGCAAGATGACGGGAAACAGAGAAATCATGAGATGTA 3184

RESULT 12
US-10-261-482-3
; Sequence 3, Application US/10261482
; Publication No. US20030036089A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: CLO00869CON
; CURRENT APPLICATION NUMBER: US/10/261,482
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 09/684,393
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: 60/172,600
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 8001
; TYPE: DNA
; ORGANISM: Human
US-10-261-482-3

Query Match 76.1%; Score 1947.4; DB 9; Length 8001;
Best Local Similarity 82.0%; Pred. No. 0;
Matches 2558; Conservative 0; Mismatches 1; Indels 560; Gaps 5;

OY 1 ATGCTGGGCGGCTGTCTCTGTGGGCTCAAGCTCTGTGGGCTCTCTGTGACCTGGAGGAGG 60
|||
DB 2118 ATGCTGGGCGGCTGTCTCTGTGGGCTCAAGCTCTGTGGGCTCTCTGTGACCTGGAGGAGG 2177

OY 61 GCGGCGATGTGCTGTGACAGCACTTAGATGAAGGGGAGCTACGTGTGGGGGGGCTG 120
|||
DB 2178 GCGGCGATGTGCTGTGACAGCACTTAGATGAAGGGGAGCTACGTGTGGGGGGGCTG 2237
OY 121 TTCGCCCTGGGAGGCGGAGAGGCTGTGCTCCGAGCGGAGCAAGGCCCGACGACCTT 180
|||
DB 2238 TTCGCCCTGGGAGGCGGAGAGGCTGTGCTCCGAGCGGAGCAAGGCCCGACGACCTT 2297
OY 181 GTGTGCA----- 187
|||
DB 2298 GTGTGCAAGGATGACAGAGTGGAGCGGCTGTGGGTGTCAGGGTACAGGTCTGGG 2357
OY 188 -----CCAGGTTCTCT 199
|||
DB 2358 GTGTCTGTGAGCTGGGCGCGAGAGTGGCACTGTGGCTTGTGTGTGGCCCGGAGTCTCT 2417
OY 200 CAAGGCGCTGTCTGTGGGCACTGTGGCAATGAATGCGCTGTGAGAGATCAACAACT 259
|||
DB 2418 CAAGGCGCTGTCTGTGGGCACTGTGGCAATGAATGCGCTGTGAGAGATCAACAACT 2477
OY 260 CGGATCTGTGCTGCGGCGCTGTGGCTGTGAGCTCTTTGATATGCTGTGGAGCTG 319
|||
DB 2478 CGGATCTGTGCTGCGGCGCTGTGGCTGTGAGCTCTTTGATATGCTGTGGAGCTG 2537
OY 320 TGGTGGCCATGAAGCGGAGCTGTGCTGTGGCCAAAGCGAGGAGCGCGGAGATGCGCG 379
|||
DB 2538 TGGTGGCCATGAAGCGGAGCTGTGCTGTGGCCAAAGCGAGGAGCGCGGAGATGCGCG 2597
OY 380 CCTACTGCACTACAGCAGTACAGCAGCCCGTGTGTGTGCTGTGATCATGCGGCGCCACTGT 439
|||
DB 2598 CCTACTGCACTACAGCAGTACAGCAGCCCGTGTGTGTGCTGTGATCATGCGGCGCCACTGT 2657
OY 440 CAGAGCTGCGCATGTGTACAGGCGGAGAGTCTTGTAGCTTCTCTCAAGCCCC----- 489
|||
DB 2658 CAGAGCTGCGCATGTGTACAGGCGGAGAGTCTTGTAGCTTCTCTCAAGCCCC----- 489
OY 490 ----- 489
|||
DB 2718 CCGCCGACCATGACCCAGCCGCGGCGGAGCGGCGGAGCGGCGGAGCGGCGGAG 2777
OY 490 -----CAGGTACAGTACAGTGTGCTAGCATGAGCTGTGAGCGCGGAG 534
|||
DB 2778 GCGCTGTGGCGCTGTGAGTACAGTGTGCTAGCATGAGCTGTGAGCGCGGAG 2837
OY 535 ACCTTCCCTCTCTTTCGCGACCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 594
|||
DB 2838 ACCTTCCCTCTCTTTCGCGACCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 2897
OY 595 GAGCTGTGCAAGAGTGTGGTGAACCTGGTGGCGCGGCGGAGCGGAGCGGAGCGGAGTAC 654
|||
DB 2898 GAGCTGTGCAAGAGTGTGGTGAACCTGGTGGCGCGGCGGAGCGGAGCGGAGCGGAGTAC 2957
OY 655 GCGGCGGAGGCGCTGTGAGATCTTTCGCGGCGGAGCGGAGCGGAGCGGAGCGGAG 714
|||
DB 2958 GCGGCGGAGGCGCTGTGAGATCTTTCGCGGCGGAGCGGAGCGGAGCGGAGCGGAG 3017
OY 715 CAGGAGGCGCTGTGGCGGCTGTGGCGGAGGAGTGTGGCGGAGGAGGAGGAGGAG 774
|||
DB 3018 CAGGAGGCGCTGTGGCGGCTGTGGCGGAGGAGTGTGGCGGAGGAGGAGGAGGAG 3077
OY 775 GTCTGTGACACAGTGAACAGAGAGCTGTGAGGTGTGTCTGTCTGTCTGTCTGTCTGT 834
|||
DB 3078 GTCTGTGACACAGTGAACAGAGAGCTGTGAGGTGTGTCTGTCTGTCTGTCTGTCTGT 3137
OY 835 GCGGCGGAGCGGCTTCTTCACTACAGCATACAGAGGAGCTGTGCGGCAAGGTGTGGGTG 894
|||
DB 3138 GCGGCGGAGCGGCTTCTTCACTACAGCATACAGAGGAGCTGTGCGGCAAGGTGTGGGTG 3197
OY 895 GCGAGCGAGGCGTGTGTGACCTGTGACCTGTGTGACCTGTGTGACCTGTGTGACCTGT 954
|||
DB 3198 GCGAGCGAGGCGTGTGTGACCTGTGTGACCTGTGTGACCTGTGTGACCTGTGTGACCTGT 3257

OY	955	GGCACGGGCGCTGGCTTCCTCCACAGAGGGGGGCCACACTGCACAGATCCCCACAGACTG	101.4
Db	3258	GGCACGGGCGCTGGCTTCCTCCACAGAGGGGGGCCACACTGCACAGATTCCTCCACAGACTG	331.7
OY	1015	AAGACGACCTGGGCGCTGGCCACCGACCGGCGCTTCGCTGCGCTGGCGAGAGGGAG	107.4
Db	3318	AAGACGACCTGGGCGCTGGCCACCGACCGGCGCTTCGCTGCGCTGGCGAGAGGGAG	337.7
OY	1075	CAGGCTCGAGAGGAGACGATGGGGGCGACGCGTGGCCGACATGATGATCAGTCAGCGTG	113.4
Db	3378	CAGGCTCGAGAGGAGACGATGGGGGCGACGCGTGGCCGACATGATGATCAGTCAGCGTG	343.7
OY	1135	CAGACGCTGAGCGAGCGGCTTAATTCACACACAGACCTTCTGCTAGCAGACTGTAT	119.4
Db	3438	CAGACGCTGAGCGAGCGGCTTAATTCACACACAGACCTTCTGCTAGCAGACTGTAT	349.7
OY	1195	AGCGTGGCCCCAGGCGCTTCGACAAACACTCTTCAGTGCACGCTCAGGCTGCCGCGCAG	125.4
Db	3498	AGCGTGGCCCCAGGCGCTTCGACAAACACTCTTCAGTGCACGCTCAGGCTGCCGCGCAG	355.7
OY	1255	GACCCCGTGAAGCGCCTG-----	127.1
Db	3558	GACCCCGTGAAGCGCCTTGGCAGATGAGCCCGGAGATGGGGGTGTGCTGTCTGTCAATGT	361.7
OY	1272	-----	127.1
Db	3618	GCCCCAGGCCAACAGGCGGCACACAGCGCTGAGCTGAGGTGGCGGCTCAGCCCC	367.7
OY	1272	-----GCAGCTCCGAGGAACATGTAAACCTGACCTTCACAGTGGCGGGGTG	132.0
Db	3678	GTCCCCCGCCCGACGCTCTCGAGAAACATGTAAACCTTCCACAGTGGCGGGGTG	373.7
OY	1331	CCGCTGGGCTTCGACACAGCAGCGGAACGATGGACATGAGTACGACCTGAAGCTGTGGTG	138.0
Db	3738	CCGCTGGGCTTCGACACAGCAGCGGAACGATGGACATGAGTACGACCTGAAGCTGTGGTG	379.7
OY	1381	TGGGAGGGCTTCAGTCCCGCAGGCTTCACAGATGGGAGGTAAACGCGACTCAGAGCA	144.0
Db	3798	TGGGAGGGCTTCAGTCCCGCAGGCTTCACAGATGGGAGGTAAACGCGACTCAGAGCA	385.7
OY	1441	GAGCGCTGAAGATCCGCTGGCACAGCTGTGACAC-----	147.6
Db	3858	GAGCGCTGAAGATCCGCTGGCACAGCTGTGACACACAGTGAAGGTGAGGGGTGTGC	391.7
OY	1477	-----	147.6
Db	3918	CAGCGTGCCTGTGTAGCCCCCGCGCAGAGGGCGACGCTGGGGGTGGGCGCTTCACG	397.7
OY	1477	-----CAGAAAGCCCG	148.6
Db	3978	TCCTCCCGTGGGCAATGCCCAACCCAGACAGACCAAGACCCAGGCTGTGCGCAGAAAGCCCG	403.7
OY	1487	TGTCGCGGTGCTCGCGGCACTGCCAGAGAGGGCGCAGGTGCGCGGGTCAAGGGGTTCACT	154.6
Db	4038	TGTCGCGGTGCTCGCGGCACTGCCAGAGAGGGCGCAGGTGCGCGGGTCAAGGGGTTCACT	409.7
OY	1547	CCCTGCTCTACGACTGTGTGGACTGTGGAGGGGGGCAAGCTAACCGGAAANCC-----	159.7
Db	4098	CCCTGCTCTACGACTGTGTGGACTGTGGAGGGGGGCAAGCTAACCGGAAANCCAGGTGAGC	415.7
OY	1598	-----	159.7
Db	4158	CGCCTTCCCGCAGCGCGGGGTGGAAACGACAGAGGGAGGGTCTCTGCCAATGCTGACT	421.7
OY	1598	-----CAG	160.0
Db	4218	CTGAGACCAAGGCCACAGGGGCAAGAGCAACCCACGCGCTTCCTCTCTCAAG	427.7
OY	1601	ACGACATCGCTGCACCTTTTGTGGCAGATGAGTGGTCCCGGAGGGAAGCACACGCT	166.0
Db	4278	ACGACATCGCTGCACCTTTTGTGGCAGATGAGTGGTCCCGGAGGGAAGCACACGCT	433.7
OY	1661	GCCTTCGCGCGAGGTCTCGGTTCTCTGGCATGGGGCGAGCGGCGTGTCTGTCTGTCTCC	172.0

Db	4338	GCCTCCGCCGAGGCTCTCCGGTCTCGGCAATGGGGCCAGCCGGCTGTGCTGCTGCTCC	4397
Qy	1721	TGCTGTAGAGCTTGCGCCTGGGCGCTTGCTGCTGCTCTTTGGGGGCTGTTGTTCAACATC	1780
Db	4398	TGCTGTAGAGCTTGCGCCTGGGCGCTTGCTGCTGCTCTTTGGGGGCTGTTGTTCAACATC	4457
Qy	1781	GGGACAGCCCACTGGCTTCAGGCGCTCGGGGGGGGGCCCTGGCGTGGCTTTGGGCTGGGTGGC	1841
Db	4458	GGGACAGCCCACTGGCTTCAGGCGCTCGGGGGGGGGCCCTGGCGTGGCTTTGGGCTGGGTGGC	4517
Qy	1841	TGGGCTGTGTTGCTTCAGGCTCCCTCTCTGTTTCTTCCCTGGCCAGCCAGCCCTGCCGATGC	1900
Db	4518	TGGGCTGTGTTGCTTCAGGCTCCCTCTCTGTTTCTTCCCTGGCCAGCCAGCCCTGCCGATGC	4577
Qy	1901	TGGCCCAAGACCCCTTGTCCCACTCCCGCTACAGGGCTGCTGAGACATCTTCTTCG	1966
Db	4578	TGGCCCAAGACCCCTTGTCCCACTCCCGCTACAGGGCTGCTGAGACATCTTCTTCG	4637
Qy	1961	AGGGGCGCGAGATTTTGTTGAGAGCAGAAATGCTCTGAGCTGGGCGACAGCCGCTGAGTG	2022
Db	4638	AGGGGCGCGAGATTTTGTTGAGAGCAGAAATGCTCTGAGCTGGGCGACAGCCGCTGAGTG	4697
Qy	2021	GCTGCTGCGGGGGCCCTGGGCGCTGGCTGGTGTCTGCTGGCATGCTGTGTGGAGATGC	2088
Db	4698	GCTGCTGCGGGGGCCCTGGGCGCTGGCTGGTGTCTGCTGGCATGCTGTGTGGAGATGC	4757
Qy	2081	CACGTGCACCTTGTTACTCTGGTGGCCCTTCCCGCGAGGGAGGGAGAGGACTGGACATGC	2140
Db	4758	CACGTGCACCTTGTTACTCTGGTGGCCCTTCCCGCGAGGGAGGGAGAGGACTGGACATGC	4817
Qy	2141	TGCCACGAGAGCGCTGGTGCACCTGCACGACAGCTCTGGGTCACTTCGGGCTAGCGC	2200
Db	4818	TGCCACGAGAGCGCTGGTGCACCTGCACGACAGCTCTGGGTCACTTCGGGCTAGCGC	4877
Qy	2201	ACGCCACCAATGCCAGCGCTGGCTTTTCTGTGCTTCTCGGGCACTTTCCTGGTGGGAGCC	2266
Db	4878	ACGCCACCAATGCCAGCGCTGGCTTTTCTGTGCTTCTCGGGCACTTTCCTGGTGGGAGCC	4937
Qy	2261	AGCGCGGCTGTCTACACCGTGGCCGTTGCGCTCACTTTCGCATGCTGGGCTCACTTCATCA	2320
Db	4938	AGCGCGGCTGTCTACACCGTGGCCGTTGCGCTCACTTTCGCATGCTGGGCTCACTTCATCA	4997
Qy	2321	CGTGGGTCTCTTTGTGCTCCCTCTCGGCCAATGTGAGTGGTCTCTCAGGCCCGCGCTGC	2388
Db	4998	CGTGGGTCTCTTTGTGCTCCCTCTCGGCCAATGTGAGTGGTCTCTCAGGCCCGCGCTGC	5057
Qy	2381	AGATGGGCGCCCTCTGCTCTGTGTGCTCGTGGGCAATTCGTGGCTGGCTTCACCTGCCAGAT	2440
Db	5058	AGATGGGCGCCCTCTGCTCTGTGTGCTCGTGGGCAATTCGTGGCTGGCTTCACCTGCCAGAT	5117
Qy	2441	GTTACTGTCTCATTGCGGCGACCGAGGGCTCAACACCCCGAGTTCTTCTGGGAGGGGCGC	2500
Db	5118	GTTACTGTCTCATTGCGGCGACCGAGGGCTCAACACCCCGAGTTCTTCTGGGAGGGGCGC	5177
Qy	2501	CTGGGAGATGCCCAAGGCGCAATAGACGGGAAACAGAGAAATCAGGGGAAACATGATGA	2559
Db	5178	CTGGGAGATGCCCAAGGCGCAATAGACGGGAAACAGAGAAATCAGGGGAAACATGATGA	5236

```

; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/187,546
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: 60/195,536
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/209,840
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: 60/214,213
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/226,448
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/259,227
; PRIOR FILING DATE: 2001-01-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 3563
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-799-629-20

```

```

Query Match      74.1%; Score 1897; DB 9; Length 3563;
Best Local Similarity 80.7%; Pred. No. 0;
Matches 2559; Conservative 0; Mismatches 0; Indels 612; Gaps 5;

```

```

QY 1 ATGCTGGGCGCTGTCTCTGAGGCTCAGCCCTGCGGCTCTCTGACCCCTGGAGCGGG 60
DB 211 ATGCTGGGCGCTGTCTCTGAGGCTCAGCCCTGCGGCTCTCTGACCCCTGGAGCGGG 270
QY 61 GCCCATTTGTGCTGTCAAGCACTTAGATGAAGGGGAGTACGTGTGGGGGGCTG 120
DB 271 GCCCATTTGTGCTGTCAAGCACTTAGATGAAGGGGAGTACGTGTGGGGGGCTG 330
QY 121 TTCCCTTGGGGGAGGCGGAGAGGCTGGCTCCGACGCGGACAGCGCCCAAGCCCT 180
DB 331 TTCCCTTGGGGGAGGCGGAGAGGCTGGCTCCGACGCGGACAGCGCCCAAGCCCT 390
QY 181 GTGTGCA----- 187
DB 391 GTGTGCAAGGTTACAGAGGTGGAGCGGCTGGGGGTGAGGTTACAGAGGTTGCGG 450
QY 188 -----CGAGTTCTCT 199
DB 451 GTGTCTGTAGCTGGGGCGGAGGTGGCATCTGCGGTTCTGTGTGGCCCAAGTTCTCT 510
QY 200 CAAAGGGCTGTCTGGGCGACTGGCCATGAATAATGGCGTGGAGAGATCAACAACAGT 259
DB 511 CAAAGGGCTGTCTGGGCGACTGGCCATGAATAATGGCGTGGAGAGATCAACAACAGT 570
QY 260 CGGATCTGTGCGCGGGGCTGGGCTGGGGTACGACCTCTTTGATACGTGTGGAGCCCTG 319
DB 571 CGGATCTGTGCGCGGGGCTGGGCTGGGGTACGACCTCTTTGATACGTGTGGAGCCCTG 630
QY 330 TGGTGGCCATGAAGCCAGCCCTCATGTTCTTGGCCAGGCGAGGCGAGCCGATCGCG 379
DB 631 TGGTGGCCATGAAGCCAGCCCTCATGTTCTTGGCCAGGCGAGGCGAGCCGATCGCG 690
QY 380 CCTATGCACTACAGAGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 439
DB 691 CCTATGCACTACAGAGATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 750
QY 440 CAGAGCTGGCCATGTGACAGGCAAGTCTTCAAGCTTCTCTCAATGAGC----- 489
DB 751 CAGAGCTGGCCATGTGACAGGCAAGTCTTCAAGCTTCTCTCAATGAGC----- 810
QY 490 ----- 489
DB 811 CCCCCCACCATTACCCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 870
QY 490 ----- 489
DB 871 GAATGCTACATGACCCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 930

```

```

QY 490 -----CAGGTACGTAAGGCTAGCATGAGAGGTGCTGAGCGCCGGGAGACCTTCC 542
DB 931 GGCCTTGAGGTGACGTACGAGGTGTGATGAGAGGTGCTGAGAGCGCCGGGAGACCTTCC 990
QY 543 CTCTCTTTCGCGACCGAGGCGGAGGAGGCTGTGAGTACAGGCGCGCGGAGCTGTCT 602
DB 991 CTCTCTTTCGCGACCGAGGCGGAGGAGGCTGTGAGTACAGGCGCGCGGAGCTGTCT 1050
QY 603 GCAGAGTTCGGCTGGAATGAGGTGCGCGCTTGGGCGACGACGACGAGTACGCGCGCA 662
DB 1051 GCAGAGTTCGGCTGGAATGAGGTGCGCGCTTGGGCGACGACGACGAGTACGCGCGCA 1110
QY 663 GGGCGTGAACATCTCTGCGGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 722
DB 1111 GGGCGTGAACATCTCTGCGGCGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1170
QY 723 CTTGTGCGCGCTGCGCGGCTGCGGATGACTCGCGGCTGGGGAGGAGTACGAGTCTGCA 782
DB 1171 CTTGTGCGCGCTGCGCGGCTGCGGATGACTCGCGGCTGGGGAGGAGTACGAGTCTGCA 1230
QY 783 CCAGTGAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 842
DB 1231 CCAGTGAACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1290
QY 843 CGCCCTCTTCACTACAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 902
DB 1291 CGCCCTCTTCACTACAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1350
QY 903 GGCCTGAGCTGACCTGTGACCTGTGATGAGGAGGCTGCGCGGAGAGGAGAGAGAGAG 962
DB 1351 GGCCTGAGCTGACCTGTGACCTGTGATGAGGAGGCTGCGCGGAGAGGAGAGAGAGAG 1410
QY 963 GCTTGGCTTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1022
DB 1411 GCTTGGCTTCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1470
QY 1023 CTTGAGCCCTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1082
DB 1471 CTTGAGCCCTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1530
QY 1083 GGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1142
DB 1531 GGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1590
QY 1143 GAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1202
DB 1591 GAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1650
QY 1203 CCAGGCGCTGCAACAACACTCTTCAAGTGAACAGCTGAGGCTGCCCGGCGAGAGCCCT 1262
DB 1651 CCAGGCGCTGCAACAACACTCTTCAAGTGAACAGCTGAGGCTGCCCGGCGAGAGCCCT 1710
QY 1263 GAAGCCCTG----- 1271
DB 1711 GAAGCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1770
QY 1272 ----- 1271
DB 1771 CACCAAGGAGGCGCACACAGCCCTGAGCTGAGAGTGGAGTGGCGCTGACGCCCTGCCG 1830
QY 1272 ---GAGCTCTGTGAGAGATGTACAACTGTGACCTTCCAGTGGGCGGCTGCCGTGG 1328
DB 1831 CCGGAGAGCTCTGTGAGAGATGTACAACTGTGACCTTCCAGTGGGCGGCTGCCGTGG 1890
QY 1329 GTTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1388
DB 1891 GTTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1950
QY 1389 CTGAGTGGCCAGAGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1448
DB 1951 CTGAGTGGCCAGAGGCTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2010
QY 1449 GAAGATCCGCTGGCAGACGCTGAGAAC----- 1476

```

```

Db 2011 GAAGATCCGCTGGACACAGCTCTGACAAACAGAGTGAAGGTGGGTGTCACAGCGGTG 2070
QY 1477 ----- 1476
Db 2071 CCCGTGTAGCCCCCGGAGAGGGGAGCCCTGGGGGTGGGGGCGTTCCAGTCTCCGT 2130
QY 1477 -----CAGAAAGCCGCTGCCG 1494
Db 2131 GGGATGCCACAGCCGAGACAGAGCCAGACCCAGGCTGTGTGGCAGAAAGCCGTTCCCG 2190
QY 1495 TGTCTGGGAGAGTCCAGAGAGGGGAGAGTGGCGCGGTCAAGGGGTTCACCTCTGTCTG 1554
Db 2191 TGTCTGGGAGAGTCCAGAGAGGGGAGAGTGGCGCGGTCAAGGGGTTCACCTCTGTCTG 2250
QY 1555 TAGCAGTGTGTGAGTCCAGAGGGGAGAGTGGCGCGGTCAAGGGGTTCACCTCTGTCTG 1597
Db 2251 TAGCAGTGTGTGAGTCCAGAGGGGAGAGTGGCGCGGTCAAGGGGTTCACCTCTGTCTG 2310
QY 1598 ----- 1597
Db 2311 CGGACGCGGGGGTGGGAGACGACAGGGGAGGTCTCCGCAAGTCCGACTCTGAGAC 2370
QY 1598 -----CAGACGACATC 1608
Db 2371 AGAGCCACAGGGTACAAGACAGAACCCAGGCCCTTCTCTCTCTCTCTCTCTCTCTCTCT 2430
QY 1609 GCGTGCACCTTTTGTGGCCAGAGATGAGTGGTCCCGGAGGAGCAGACAGCGTGTCTCCG 1668
Db 2431 GCGTGCACCTTTTGTGGCCAGAGATGAGTGGTCCCGGAGGAGCAGACAGCGTGTCTCCG 2490
QY 1669 CGCAGTCTCGGTCTCTGAGATGGGGGAGAGCGGCTGTCTCTCTCTCTCTCTCTCTCTCT 1728
Db 2491 CGCAGTCTCGGTCTCTGAGATGGGGGAGAGCGGCTGTCTCTCTCTCTCTCTCTCTCTCT 2550
QY 1729 AGCCTGGCGCTGGGCGCTTGTGTCTGTCTGTGGGCGTGTCTCTCTCTCTCTCTCTCTCT 1788
Db 2551 AGCCTGGCGCTGGGCGCTTGTGTCTGTCTGTGGGCGTGTCTCTCTCTCTCTCTCTCTCT 2610
QY 1789 CCACTGGTTCAGGCTGGGGGAGGCGCCCTGAGCTTGTGGCGTGGTGGTGGTGGTGGTGGT 1848
Db 2611 CCACTGGTTCAGGCTGGGGGAGGCGCCCTGAGCTTGTGGCGTGGTGGTGGTGGTGGTGGT 2670
QY 1849 GTCTGCTCAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1908
Db 2671 GTCTGCTCAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2730
QY 1909 CAGGCTGTTCACCTCCGCTCAGAGGCGCTGAGCAGACACTCTCTCTCTCTCTCTCTCTCT 1968
Db 2731 CAGGCTGTTCACCTCCGCTCAGAGGCGCTGAGCAGACACTCTCTCTCTCTCTCTCTCTCT 2790
QY 1969 GAGATCTTCTGAGTGAAGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2028
Db 2791 GAGATCTTCTGAGTGAAGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2850
QY 2029 CGGGGCGCTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2088
Db 2851 CGGGGCGCTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2910
QY 2089 ACCGTGATACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2148
Db 2911 ACCGTGATACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2970
QY 2149 GAGGCGCTGGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2208
Db 2971 GAGGCGCTGGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3030
QY 2209 AATGCAACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2268
Db 3031 AATGCAACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3090
QY 2269 TGTACAAACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2338

```

```

Db 3091 TGTACAAACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3150
QY 2329 TCCCTTGTGCCCCCTCCGAGCAATGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2388
Db 3151 TCCCTTGTGCCCCCTCCGAGCAATGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3210
QY 2389 GCGCTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2448
Db 3211 GCGCTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3270
QY 2449 CTCATGTGGGAGCCAGGAGCTCAACACCCCGAGTCTTCTCTGTGGAGGGGCGCTGGGAT 2508
Db 3271 CTCATGTGGGAGCCAGGAGCTCAACACCCCGAGTCTTCTCTGTGGAGGGGCGCTGGGAT 3330
QY 2509 GCCCAAGGCCAGATGACGGGAGACAGAAATCAGGGGAAACATGATGA 2559
Db 3331 GCCCAAGGCCAGATGACGGGAGACAGAAATCAGGGGAAACATGATGA 3381

```

```

RESULT 14
US-10-035-045-20
; Sequence 20, Application US/10035045
; Publication No. US2003005448A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIADONG
; APPLICANT: STASZEMSKI, LENA
; APPLICANT: O'CONNELL, SHAWN
; APPLICANT: ZOZULYA, SERGEY
; TITLE OR INVENTION: TIR TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003-0280681
; CURRENT APPLICATION NUMBER: US/10/035,045
; PRIOR FILING DATE: 2002-01-03
; PRIOR APPLICATION NUMBER: 60/259,227
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: 60/284,547
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 20
; LENGTH: 3563
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-035-045-20

```

```

Query Match 74.1%; Score 1897; DB 9; Length 3563;
Best Local Similarity 80.7%; Pred. No. 0;
Matches 2559; Conservative 0; Mismatches 0; Indels 612; Gaps 5;

```

```

QY 1 ATGCTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60
Db 211 ATGCTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 270
QY 61 GCCCATTTGTCTCTGTACAGCACTTGAAGTGAAGGGGAGCTGAGTGGGGGGGCTG 120
Db 271 GCCCATTTGTCTCTGTACAGCACTTGAAGTGAAGGGGAGCTGAGTGGGGGGGCTG 330
QY 121 TTCCCGCTGGGCGAGGCGAGAGGAGTGGCTCCGACAGCCGAGACCGGAGAGCCCT 180
Db 331 TTCCCGCTGGGCGAGGCGAGAGGAGTGGCTCCGACAGCCGAGACCGGAGAGCCCT 390
QY 181 GTGTGCA----- 187
Db 391 GTGTGCAACAGGTACAGAGGTGGAGCGGCTGGGTGGGGGTGACAGGTCTGGG 450
QY 188 -----CAGGTTCTCT 199
Db 451 GTGTCTCTGAGTGGGGCGAGAGTGGCCATCTGCGGTTCTGTGTGTGTGTGTGTGTGT 510
QY 200 CAACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 259
Db 511 CAACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 570

```

260 CGATCTGCTGCGCGGCTGCGCTGAGGTACAGACTCTTTGATAGTGTCTCGAGACCTG 319
|||||
Db 571 CGGATCTGTGTGCGCGGCTGCGCTGAGGTACAGACTCTTTGATAGTGTCTCGAGACCTG 630
|||||
Oy 320 TGTGTGCGCATGAGGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCATGCGCG 379
|||||
Db 631 TGTGTGCGCATGAGGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCATGCGCG 690
|||||
Oy 380 CTTACTGCACTACAGCGAGTACAGCGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 439
|||||
Db 691 CTTACTGCACTACAGCGAGTACAGCGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 750
|||||
Oy 440 CAGAGCTGCGCATGAGTACAGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 489
|||||
Db 751 CAGAGCTGCGCATGAGTACAGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 810
|||||
Oy 490 ----- 489
Db 811 CCGCCACCATGACCCCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 870
|||||
Oy 490 ----- 489
Db 871 GAATGCTACATGACCCCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 930
|||||
Oy 490 ----- 489
Db 931 GCGCTTGCAGAGTACAGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 990
|||||
Oy 543 CTTCTTCTTCCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 602
|||||
Db 991 CTTCTTCTTCCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1050
|||||
Oy 603 GCGAGAGTTCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 662
|||||
Db 1031 GCGAGAGTTCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1110
|||||
Oy 663 GCGCGTGCAGATCTTCTGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 722
|||||
Db 1111 GCGCGTGCAGATCTTCTGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1170
|||||
Oy 723 CTTGTGCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 782
|||||
Db 1171 CTTGTGCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1230
|||||
Oy 783 CCGAGTGAACGAGAGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 842
|||||
Db 1231 CCGAGTGAACGAGAGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1290
|||||
Oy 843 CCGCGTTCCTCACTACAGATGAGAGCGAGCGCGCATGATGTTCTGTGCGCAAGGACGAGCG 902
|||||
Db 1291 CCGCGTTCCTCACTACAGATGAGAGCGAGCGCGCATGATGTTCTGTGCGCAAGGACGAGCG 1350
|||||
Oy 903 GCGCTGCGTGCAGCTGATGAGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 962
|||||
Db 1351 GCGCTGCGTGCAGCTGATGAGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1410
|||||
Oy 963 GCTTGGCTTCCTCAAGAGGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1022
|||||
Db 1411 GCTTGGCTTCCTCAAGAGGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1470
|||||
Oy 1023 CTTGTGCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1082
|||||
Db 1471 CTTGTGCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1530
|||||
Oy 1083 GCGAGAGAGAGT 1142
|||||
Db 1531 GCGAGAGAGAGT 1590
|||||
Oy 1143 GAGCGAGAGAGT 1202
|||||
Db 1591 GAGCGAGAGAGT 1650
|||||
Oy 1203 CAGCGCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1262
|||||

1651 CGAGCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1710
|||||
Oy 1263 GAGCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1271
|||||
Db 1711 GAGCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1770
|||||
Oy 1272 ----- 1271
Db 1771 CAGCAGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1830
|||||
Oy 1272 ----- 1271
Db 1831 CCGCGAGAGT 1890
|||||
Oy 1329 GTTTGCAGCAGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1388
|||||
Db 1891 GTTTGCAGCAGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1950
|||||
Oy 1389 CTTAGTGCAGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1448
|||||
Db 1951 CTTAGTGCAGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 2010
|||||
Oy 1449 GAGAGTCCGCTGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1476
|||||
Db 2011 GAGAGTCCGCTGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 2070
|||||
Oy 1477 ----- 1476
Db 2071 CCGGTGTAGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 2130
|||||
Oy 1477 ----- 1476
Db 2131 GCGCATGCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 2190
|||||
Oy 1495 TGTGTGCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1554
|||||
Db 2191 TGTGTGCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 2250
|||||
Oy 1555 TACGACTGTGTGAGTGTGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1597
|||||
Db 2251 TACGACTGTGTGAGTGTGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 2310
|||||
Oy 1598 ----- 1597
Db 2311 CCGCAGCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 2370
|||||
Oy 1598 ----- 1597
Db 2371 AGAGCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 2430
|||||
Oy 1603 GCGTGCAGCTTGT 1668
|||||
Db 2431 GCGTGCAGCTTGT 2490
|||||
Oy 1669 CCGAGGTTCGCGT 1728
|||||
Db 2491 CCGAGGTTCGCGT 2550
|||||
Oy 1729 AGCGTGCAGCTTGT 1788
|||||
Db 2551 AGCGTGCAGCTTGT 2610
|||||
Oy 1789 CCACTGT 1848
|||||
Db 2611 CCACTGT 2670
|||||
Oy 1849 GTCGTGCAGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1908
|||||
Db 2671 GTCGTGCAGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 2730
|||||
Oy 1909 CAGCGCGCGCGCGCGCATGATGTTCTGTGCGCAAGGACGAGCGCGCAT 1968
|||||


```

Db 2731 CAGCCCTTGTCCACCTCCGCTACAGGGCTGCTGTAGACACACTCTTCTCAGAGCGGCC 2790
QY 1969 GAGATCTTCTGAGTGAAGAACTGCTCTGAGTGGGAGAGACGGGCTGAGTGGCTGTG 2028
Db 2791 GAGATCTTCTGAGTGAAGAACTGCTCTGAGTGGGAGAGACGGGCTGAGTGGCTGTG 2850
QY 2029 CGGGGGCCCTGGGCTGAGTGGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2088
Db 2851 CGGGGGCCCTGGGCTGAGTGGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2910
QY 2089 ACCTGTACCTGAGTGGGCTTCCCGCGAGGTGGTGAAGAGGATGGGATGCTGCTGCTGCTG 2148
Db 2911 ACCTGTACCTGAGTGGGCTTCCCGCGAGGTGGTGAAGAGGATGGGATGCTGCTGCTGCTG 2970
QY 2149 GAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2208
Db 2971 GAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3030
QY 2209 AATGCCACGCTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2268
Db 3031 AATGCCACGCTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3090
QY 2269 TGCTACACACGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2328
Db 3091 TGCTACACACGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3150
QY 2329 TCTTTGTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2388
Db 3151 TCTTTGTGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3210
QY 2389 GCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2448
Db 3211 GCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3270
QY 2449 CTCATGGGCGAGCGAGGCTCAACACCCCGAGTTCTTCTGAGAGGGGCGCTGGGGAT 2508
Db 3271 CTCATGGGCGAGCGAGGCTCAACACCCCGAGTTCTTCTGAGAGGGGCGCTGGGGAT 3330
QY 2509 GCCCAAGGCCAAGATGACGGGAGACAGAGAAATCAGGGGAAATGAGTGA 2559
Db 3331 GCCCAAGGCCAAGATGACGGGAGACAGAGAAATCAGGGGAAATGAGTGA 3381

RESULT 15
; Sequence 2, Application US/09799629
; Publication No. US2003008344A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: ZOZULYA, SERGEY
; APPLICANT: LI, XIADONG
; APPLICANT: O'CONNELL, SHAWN
; APPLICANT: STASZEWSKI, LENA
; TITLE OF INVENTION: T1R TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003/0277870/RXT
; CURRENT APPLICATION NUMBER: US/09/799,629
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/187,546
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: 60/195,536
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/209,840
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: 60/214,213
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/226,448
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/259,227
; PRIOR FILING DATE: 2001-01-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2687

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-799-629-2

Query Match      65.1%; Score 1666; DB 9; Length 2687;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2070; Conservative 0; Mismatches 0; Indels 374; Gaps 3;

QY 490 CAGGTGACGTGAGTGTGATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 549
Db 62 CAGGTGACGTGAGTGTGATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 121
QY 550 TTCGCGACCGTGGCCAGGAGACCTGTGACGTGACGAGGCGCGGAGCTGCTGACAGAG 609
Db 122 TTCGCGACCGTGGCCAGGAGACCTGTGACGTGACGAGGCGCGGAGCTGCTGACAGAG 181
QY 610 TTGGGTGGAACGTGGGTGGGCGCCCTGGGCGACAGCAGCAGTACGGGCGGAGAGGGCTG 669
Db 182 TTGGGTGGAACGTGGGTGGGCGCCCTGGGCGACAGCAGCAGTACGGGCGGAGAGGGCTG 241
QY 670 AGCATCTTCTCGGCCCTGGGCGGCGCAAGCGGCACTGTGATCGGCGCAAGAGGCGCTG 729
Db 242 AGCATCTTCTCGGCCCTGGGCGGCGCAAGCGGCACTGTGATCGGCGCAAGAGGCGCTG 301
QY 730 CCGCTGCCCGCTGCCGATGACTGCGCGCTGGGGAAGTGTGACAGACGTCCTGACAGAGTG 789
Db 302 CCGCTGCCCGCTGCCGATGACTGCGCGCTGGGGAAGTGTGACAGACGTCCTGACAGAGTG 361
QY 790 AACCAGACACGCTGACAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 849
Db 362 AACCAGACACGCTGACAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 421
QY 850 TTCACTACACACATCAGACAGAGGCTCTGCGCCAAAGTGTGGTGGCGACGAGGCGCTG 909
Db 422 TTCACTACACACATCAGACAGAGGCTCTGCGCCAAAGTGTGGTGGCGACGAGGCGCTG 481
QY 910 CTGACCTGTGACCTGTGATGATGAGGCTGCGCGGATGAGGCGGAGTGGGCTGCTGAGC 969
Db 482 CTGACCTGTGACCTGTGATGATGAGGCTGCGCGGATGAGGCGGAGTGGGCTGCTGAGC 541
QY 970 TTCTCCAGAGAGGGTGGCCACAGCTGACAGAGTTCCCGAGTACGTGAAGACACACTGAGC 1029
Db 542 TTCTCCAGAGAGGGTGGCCACAGCTGACAGAGTTCCCGAGTACGTGAAGACACACTGAGC 601
QY 1030 CTGGCCACCGACCCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1089
Db 602 CTGGCCACCGACCCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 661
QY 1090 GACGTGTGGGCGACGCGCTGCGCCAGTGTGACTGATCAGCTGACAGAGCTGAGCGCA 1149
Db 662 GACGTGTGGGCGACGCGCTGCGCCAGTGTGACTGATCAGCTGACAGAGCTGAGCGCA 721
QY 1150 GGGCTAATATCACACACAGAGCTTCTGTGTACAGAGTGTGTATGAGTGGGCGCCAGGC 1209
Db 722 GGGCTAATATCACACACAGAGCTTCTGTGTACAGAGTGTGTATGAGTGGGCGCCAGGC 781
QY 1210 CTGCAACAACCTTTCAGTGTGACAGGCTGACGCTGCGCGGAGAGAGAGAGAGAGAGAG 1269
Db 782 CTGCAACAACCTTTCAGTGTGACAGGCTGACGCTGCGCGGAGAGAGAGAGAGAGAGAG 841
QY 1270 TG----- 1271
Db 842 TGGCAGGTGAGCCCGGAGATGAGGGGTGTGCTGTCTCTGTGATGTGCCAGGCGACAGG 901
QY 1272 -----GAG 1275
Db 902 CAGGCGCAACACGCTGAGCTGGAGGTGGCTGGCGGCTCAAGCCCGCTCCCGCCGCGAG 961
QY 1276 CTCCTGAGAGACATGTACAACTGACCTTCCACGTGGGCGGAGCTGCGGCTGCTGAC 1335
Db 962 CTCCTGAGAGACATGTACAACTGACCTTCCACGTGGGCGGAGCTGCGGCTGCTGAC 1021
QY 1336 AGCAGCGGAAACGTGACATGTGAGTACGACCTGAAGCTGTGGGTGTGGCAGAGGCTCAGTG 1395

```



```

Db 1022 AGCAGCGAAAGCTGGACATGAGATACGACCTGAAAGCTGTGGGTGGAGGGCTCAGG 1081
OY 1386 CCCAGGCTCCACGACGTGGGGAGGTTCAAAGGAGCTTCAGAGACAGAGCCCTGAAGATC 1455
Db 1082 CCCAGGCTCCACGACGTGGGGAGGTTCAAAGGAGCTTCAGAGACAGAGCCCTGAAGATC 1441
OY 1456 CGCTGGACACGCTTGACAAAC----- 1476
Db 1142 CGCTGGACACGCTTGACAAACAGGTGAGGTGAGGTGGGTGCCAGGCGTGGCCGTGG 1201
OY 1477 ----- 1476
Db 1202 TAGCCCCCGGCGAGGGCCGACACCTGGGGGTGGGGCCCTTCACATCTCCCTGGGCAATG 1261
OY 1477 -----CAGAGCGCGGTGTCCCGGCTCGC 1501
Db 1262 CCCAGCCGAGACGACGACAGACCCAGGCTGTGCGCAGAAAGCCGTTGCCGTGCTCGC 1321
OY 1502 GGCAGTGGCAGGAGGGCCAGGTGGCGCGGTTCAAGGGGTCCACCTGCTGCTCAAGACT 1561
Db 1322 GGCAGTGGCAGGAGGGCCAGGTGGCGCGGTTCAAGGGGTCCACCTGCTGCTCAAGACT 1381
OY 1562 GTGTGACTGCGAGCGGGGACACTACCGCAAAACC----- 1597
Db 1382 GTGTGACTGCGAGCGGGGACACTACCGCAAAACCAGGTGAGCGGCTTCGCGGCAAG 1441
OY 1598 ----- 1597
Db 1442 CGGGGGTGGGAACGAGCAGAGGGGAGGTCTGCAAGTCTGACTGTGAGCAGAGCC 1501
OY 1598 -----CAGACGACATGGCTGCA 1615
Db 1502 ACAGGTTACAGACGACACCCAGGCGCTTCTCTCTCTCACAGACGACATGCTGCA 1561
OY 1616 CCTTTTGGGCCAGGATGATGTTCCCGGAGCGAAGCAGACGCTGCTTCGCGCAGGT 1675
Db 1562 CCTTTTGGGCCAGGATGATGTTCCCGGAGCGAAGCAGACGCTGCTTCGCGCAGGT 1621
OY 1676 CTGCGTCTGCGATGGGGCGAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1735
Db 1622 CTGCGTCTGCGATGGGGCGAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1681
OY 1736 CGCTGGGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1795
Db 1682 CGCTGGGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1741
OY 1796 TTTCAAGGCTCGGGGGGCGCTGAGCTGCTTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1855
Db 1742 TTTCAAGGCTCGGGGGGCGCTGAGCTGCTTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1801
OY 1856 TCAGGCTCTCTGTTCCCTGCGCAGCCAGCCCTGCGGATGCTGTGCGCAGAGCCCT 1915
Db 1802 TCAGGCTCTCTGTTCCCTGCGCAGCCAGCCCTGCGGATGCTGTGCGCAGAGCCCT 1861
OY 1916 TGTCCACCTCCCGCTCAAGGCTGCTGAGACACTTCTCTGACAGGCGCGGAGATCT 1975
Db 1862 TGTCCACCTCCCGCTCAAGGCTGCTGAGACACTTCTCTGACAGGCGCGGAGATCT 1921
OY 1976 TGTGGAAGTCAAGACTGCTGAGCTGGGCGAGACCGGCTGAGTGGCTGTGCGGGGCG 2035
Db 1922 TGTGGAAGTCAAGACTGCTGAGCTGGGCGAGACCGGCTGAGTGGCTGTGCGGGGCG 1981
OY 2036 CCTGGGCGGCTGTGGGTGCTGCTGGGCGATGCTGTGGAGGTGCGACATGTGCTGCTGCTGCT 2095
Db 1982 CCTGGGCGGCTGTGGGTGCTGCTGGGCGATGCTGTGGAGGTGCGACATGTGCTGCTGCTGCT 2041
OY 2096 ACCTGGTGGCTTCCCGCGGAGGTGCTGACGAGTGGACATGCTGCGCAGGAGGCGC 2155
Db 2042 ACCTGGTGGCTTCCCGCGGAGGTGCTGACGAGTGGACATGCTGCGCAGGAGGCGC 2101
OY 2156 TGTGTGACTGCGCGACAGGCTCTGTGGGTGAGTTCGGCTTACGCGACGCCCAATGCGCA 2215

```

```

Db 2102 TGTGTGACTGCCGCGACAGCGCTCTGGGTGACGTTGGGCTAGCGGACGCCACCAATGCGCA 2161
OY 2216 CGTGGGCTTCTGCTGTGCTGTGCGACTTTCCTGTGGGAGCGAGCCGGGCTGCTACA 2275
Db 2162 CGTGGGCTTCTGCTGTGCTGTGCGACTTTCCTGTGGGAGCGAGCCGGGCTGCTACA 2221
OY 2276 ACCGTGCCCGTGGCGCTGACGCTTTCGACATGCTGGGCTTCTGACATGACGCTGCTGCTGCTG 2335
Db 2222 ACCGTGCCCGTGGCGCTGACGCTTTCGACATGCTGGGCTTCTGACATGACGCTGCTGCTGCTG 2281
OY 2336 TGCCTCTCTGCGCAATGTCAGGTGCTCTAGGCGCGCGCTGAGATGGGCGGCTTC 2395
Db 2282 TGCCTCTCTGCGCAATGTCAGGTGCTCTAGGCGCGCGCTGAGATGGGCGGCTTC 2341
OY 2396 TGTCTGTGCTCTGCGCAATGTCAGGTGCTCTAGGCGCGCGCTGAGATGGGCGGCTTC 2455
Db 2342 TGTCTGTGCTCTGCGCAATGTCAGGTGCTCTAGGCGCGCGCTGAGATGGGCGGCTTC 2401
OY 2456 GGCAGCGAGGGCTCAAAACCCCGAGTTCTCTGSGAGGGGGCCCTGGGGATGGCCAAAG 2515
Db 2402 GGCAGCGAGGGCTCAAAACCCCGAGTTCTCTGSGAGGGGGCCCTGGGGATGGCCAAAG 2461
OY 2516 GCCAGAAATGACGGGAACACAGGAAATCAAGGGGAAACATGAGTGA 2559
Db 2462 GCCAGAAATGACGGGAACACAGGAAATCAAGGGGAAACATGAGTGA 2505

```

Search completed: May 23, 2003, 18:21:39
Job time : 399.884 secs

THIS PAGE BLANK (USPTO)

XX WO200164893-A2
 PN

Human GPCR protein
1138-rner rat protein
Rat sensory transduc
Rat sensory transduc
Rat sensory transduc
Rat sensory transduc
Mouse sensory trans
Human taste recept
Human G-protein co
Human novel G-prob
Chicken calcium-se
Human novel G-prot
Parathyroid calcul
Parathyroid calcul
Human parathyroid
Human calcium recee
Human parathyroid
Human parathyroid
Human calcium recee
Human wild type ca
Cynomolgus monkey
Protein encoded by
hNCARL.0. Homo sa
Bovine parathyroid
Bovine parathyroid
Bovine parathyroid
Bovine parathyroid
Bovine parathyroid
Bovine calcium recee
Bovine calcium recee
Parathyroid calculu
Human parathyroid
Human parathyroid
Human parathyroid
Human parathyroid
Human calcium recee

PD 07-SEP-2001.
 XX
 XX 28-FEB-2001; 2001WO-US06543.
 XX
 PR 29-FEB-2000; 2000US-186059P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Glucksmann MA, Galvin KM, Silos-Santiago I;
 XX
 DR WPI: 2001-589866/66.
 DR N-PSDB: AAS14574.
 XX
 PT Novel G protein coupled receptors and nucleic acids encoding them, for
 PT identifying agents for the treatment of cardiac disorders
 XX
 PS Claim 9; Fig 15; 209pp; English.
 XX
 CC The invention relates to novel human G protein-coupled receptors (GPCR)
 CC named 1983, 52881, 2398, 45449, 50289 and 52872. The novel GPCRs and
 CC nucleic acids encoding them are useful for identifying agents for the
 CC treatment of cardiovascular disorders, angiogenesis-related disorders,
 CC neural disorders, pain response disorders and inflammatory disorders
 CC e.g. atherosclerosis, angina pectoris and myocardial infarction,
 CC ischaemic heart disease, sudden cardiac death, hypertensive heart
 CC disease, diabetes, prostate cancer-related pain, diabetes and obesity.
 CC The present sequence represents GPCR 50289.
 CC
 XX
 SQ Sequence 852 AA;

Query Match 99.7%; Score 4512; DB 22; Length 852;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 851; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MCGPVLGLSLMALHPETGAPLCLSQQLRMKGDVYLGFLPGLGAEAGLRSTRPSP 60
 DB 1 MCGPVLGLSLMALHPETGAPLCLSQQLRMKGDVYLGFLPGLGAEAGLRSTRPSP 60
 QY 61 VCTRESSNGLMALAMKMAVEINNKSDLLPGLRGLYDLFTCPSPVYAMRPSLMLAKA 120
 DB 61 VCTRESSNGLMALAMKMAVEINNKSDLLPGLRGLYDLFTCPSPVYAMRPSLMLAKA 120
 QY 121 GSRDIAACNTQYQYPRVLAIVGPHSSSLAMVTGTFSEFSLMPQVSYGASMSLLSARETF 180
 DB 121 GSRDIAACNTQYQYPRVLAIVGPHSSSLAMVTGTFSEFSLMPQVSYGASMSLLSARETF 180
 QY 181 PSFETVSDRVOJLAAELLOEFEGMNVAALGSDDEYGRGLSTFSALAAAGCICIAHE 240
 DB 181 PSFETVSDRVOJLAAELLOEFEGMNVAALGSDDEYGRGLSTFSALAAAGCICIAHE 240
 QY 241 GLVPLPRADSRGLKVDVYLHOVNOSSVOYVLLFASVAHAHALFNYSISSRLSPKVVAS 300
 DB 241 GLVPLPRADSRGLKVDVYLHOVNOSSVOYVLLFASVAHAHALFNYSISSRLSPKVVAS 300
 QY 301 EAMLTSDLVMLPGMAQMGTVLGLQAGQIHEPPQYKTHALATDPAFCALGEREOG 360
 DB 301 EAMLTSDLVMLPGMAQMGTVLGLQAGQIHEPPQYKTHALATDPAFCALGEREOG 360
 QY 361 LEEVDVGRCPOQDCITLQNTSAGLNHQTFSYAAVSVAAQALNLTQCNASGCPADP 420
 DB 361 LEEVDVGRCPOQDCITLQNTSAGLNHQTFSYAAVSVAAQALNLTQCNASGCPADP 420
 QY 421 VKPMOLENMTNLFHFVGLPLRFDSSGNVMEYDLKLMWQSGVPRLDVGRFNGSLRT 480
 DB 421 VKPMOLENMTNLFHFVGLPLRFDSSGNVMEYDLKLMWQSGVPRLDVGRFNGSLRT 480
 QY 481 ERLKTRWTSNOKPVSRCSQCEGVYRVKGFHSCCYDCVCEAGSYRONPDDIACF 540
 DB 481 ERLKTRWTSNOKPVSRCSQCEGVYRVKGFHSCCYDCVCEAGSYRONPDDIACF 540
 QY 541 CGODEMSERSTRCFRRRSRLANGEPVALLLLLLSLATGLVLAALGLFVHHRDSPLVQ 600
 DB 541 CGODEMSERSTRCFRRRSRLANGEPVALLLLLLSLATGLVLAALGLFVHHRDSPLVQ 600

QY 601 ASGPIACFGLVCLIGVCLSVLLEPGQSPARCIAQOPLSHLPLTGCLSTLQAAELTFV 660
 DB 601 ASGPIACFGLVCLIGVCLSVLLEPGQSPARCIAQOPLSHLPLTGCLSTLQAAELTFV 660
 QY 661 ESELPISWADRLSCGLRGPAMLVLLAMLVVALCTWYLVAFPEVYTDHMLPTREALV 720
 DB 661 ESELPISWADRLSCGLRGPAMLVLLAMLVVALCTWYLVAFPEVYTDHMLPTREALV 720
 QY 721 HCTRSMWSPGLATNATNATLAFICFGLFVRSQPGCVNRARGLTPMLAVETWVSFVP 780
 DB 721 HCTRSMWSPGLATNATNATLAFICFGLFVRSQPGCVNRARGLTPMLAVETWVSFVP 780
 QY 781 LLANVOVLRPAVOMGALLICVLGILAAFLPRCYLLMRQGLMTPPEFLGGPGDAQO 840
 DB 781 LLANVOVLRPAVOMGALLICVLGILAAFLPRCYLLMRQGLMTPPEFLGGPGDAQO 840
 QY 841 NDGNTGNQKHE 852
 DB 841 NDGNTGNQKHE 852

RESULT 2
 ID AAO21501 standard; Protein; 852 AA.
 AC AAO21501;
 XX
 DT 15-AUG-2002 (first entry)
 XX
 DE Human G-protein coupled receptor (GPCR).
 XX
 KW G-protein coupled receptor; GPCR; gene chip; human; immune response;
 KW chromosomal position; transgenic animal; gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Region
 FT
 FT Region
 FT
 FT Modified-site
 FT
 FT Region
 FT
 FT Modified-site
 FT
 FT Domain
 FT
 FT Region
 FT
 FT Region
 FT
 FT Region
 FT
 FT Region
 FT
 FT Domain
 FT
 FT Modified-site
 FT
 FT Modified-site
 FT
 FT Modified-site
 FT
 FT Region

Location/Qualifiers
 20..25
 /note="Myristoylation site"
 69..74
 /note="Myristoylation site"
 85..88
 /note="Asn is N-glycosylated"
 92..97
 /note="Myristoylation site"
 102..105
 /note="Casein kinase II phosphorylation site"
 130..133
 /note="Casein kinase II phosphorylation site"
 137..167
 /note="Asn is N-glycosylated"
 137..167
 /label="Helix_1"
 /note="Membrane spanning structure and domain"
 153..155
 /note="Protein kinase C phosphorylation site"
 175..177
 /note="Protein kinase C phosphorylation site"
 175..178
 /note="Casein kinase II phosphorylation site"
 189..191
 /note="Protein kinase C phosphorylation site"
 214..217
 /note="Casein kinase II phosphorylation site"
 234..239
 /note="Myristoylation site"
 256..286
 /label="Helix_3"
 /note="Membrane spanning structure and domain"
 264..267
 /note="Asn is N-glycosylated"
 285..288
 /note="Asn is N-glycosylated"
 289..291

FT	/note="Protein kinase C phosphorylation site"
FT	293..295
FT	/note="Protein kinase C phosphorylation site"
FT	306..326
FT	/label="Helix_4
FT	319..324
FT	/note="Membrane spanning structure and domain"
FT	380..383
FT	/note="Myristoylation site"
FT	411..414
FT	/note="Asn is N-glycosylated"
FT	432..435
FT	/note="Asn is N-glycosylated"
FT	475..478
FT	/note="Asn is N-glycosylated"
FT	476..481
FT	/note="Myristoylation site"
FT	477..479
FT	/note="Protein kinase C phosphorylation site"
FT	480..482
FT	/note="Protein kinase C phosphorylation site"
FT	517..541
FT	/note="G-protein coupled receptor family 3 signature 2"
FT	528..530
FT	/note="Protein kinase C phosphorylation site"
FT	551..553
FT	/note="Protein kinase C phosphorylation site"
FT	556..559
FT	/note="cGMP-dependent protein kinase phosphorylation site"
FT	567..587
FT	/label="Helix_6
FT	581..586
FT	/note="Myristoylation site"
FT	603..608
FT	/note="Myristoylation site"
FT	606..626
FT	/label="Helix_7
FT	637..657
FT	/label="Helix_8
FT	646..651
FT	/note="Membrane spanning structure and domain"
FT	667..670
FT	/note="Casein kinase II phosphorylation site"
FT	680..700
FT	/label="Helix_9
FT	731..736
FT	/note="Membrane spanning structure and domain"
FT	731..751
FT	/label="Helix_10
FT	737..740
FT	/note="Asn is N-glycosylated"
FT	763..768
FT	/note="Myristoylation site"
FT	763..783
FT	/label="Helix_11
FT	792..812
FT	/note="Membrane spanning structure and domain"
FT	804..809
FT	/label="Helix_12
FT	831..836
FT	/note="Myristoylation site"
FT	839..844
FT	/note="Myristoylation site"
FT	843..848

FT	XX		/note= "Myristoylation site"
PN	XX		
WO	002030981-AI.		
PD	18-APR-2002.		
PF	13-MAR-2001; 2001WO-US07832.		
PR	10-OCT-2000; 2000US-0684393.		
PA	(PEKE) PE CORP NY.		
PI	Weil M, Zhong W, Ketchum KA, Difrancesco V, Beasley EM;		
DR	WPI: 2002-444173/47.		
N-	P-SDB; AAL38462, AAL38463.		
PT	Novel G protein coupled receptor, useful for raising antibodies, to		
quantitatively determine levels of protein in biological samples -			
Claim 1; Fig 2; 82pp; English.			
The invention relates to an isolated G-protein coupled receptor (GPCR)			
polypeptide, comprising an 852 residue amino acid sequence, given in the			
specification, an allelic variant or ortholog of the protein, or a			
fragment comprising at least 10 contiguous amino acids of the protein.			
GPCR is useful for identifying a modulator of GPCR and an agent that			
binds to GPCR. GPCR and a gene chip comprising GPCR are useful as models			
for the development of human therapeutic agents. GPCR is useful for			
raising antibodies, to elicit immune response, as a reagent in assays			
designed to quantitatively determine levels of protein in biological			
samples, and as markers for tissues in which the corresponding protein			
is preferentially expressed. A gene chip containing GPCR is also useful			
as a probe for determining the chromosomal positions of nucleic acid			
molecules by means of in situ hybridisation, in making vectors containing			
the gene regulatory regions of a gene chip containing GPCR, for designing			
ribozymes, in making vectors that express GPCR, and for constructing host			
cells and transgenic animals expressing nucleic acid molecules and			
peptides. A host cell containing GPCR is useful for conducting cell-based			
assays involving GPCR protein or its fragments, and for identifying			
protein mutants in which these function is affected. The polynucleotide			
encoding GPCR can be used to treat disorders by gene therapy. This			
sequence represents the human G-protein coupled receptor of the			
invention.			
Sequence 852 AA:			
Query Match	99.7%; Score 4512; DB 23; Length 852;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 851; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
1 MGPAVLGIISWALLHPTGAPLCLISQOLRMKGDIYGLFPLGEAEERGLSRTPSSP	60		
1 MGPAVLGIISWALLHPTGAPLCLISQOLRMKGDIYGLFPLGEAEERGLSRTPSSP	60		
61 VCTRSSNGLMALMKMAVEEINKSDLLPGIRLGIDLPDCCSEVYAMKRSIMFLAQA	120		
61 VCTRSSNGLMALMKMAVEEINKSDLLPGIRLGIDLPDCCSEVYAMKRSIMFLAQA	120		
121 GSRLDAAYCNYQOYPRVLAIVGPSHSSELAMYGKFSPFLMPOYSYGASMELSARETF	180		
121 GSRLDAAYCNYQOYPRVLAIVGPSHSSELAMYGKFSPFLMPOYSYGASMELSARETF	180		
181 PSFEFTVPSDRVQLTAALAELOEFGMWNVAALGSDEYGRGSLIFSALAARGICIAHE	240		
181 PSFEFTVPSDRVQLTAALAELOEFGMWNVAALGSDEYGRGSLIFSALAARGICIAHE	240		
241 GLVPFRPADDSNLGKVQDYLVHOYNOSVOYLIFLASNAHAHLFNYSISSRLSPRYWAS	300		
241 GLVPFRPADDSNLGKVQDYLVHOYNOSVOYLIFLASNAHAHLFNYSISSRLSPRYWAS	300		
301 EAWLNSDLVMSIGMAOMGTVLGFLOGSGLHPEFPQYVETHALMTDPAGCAIGFERREG	360		

Db 541 CGODEMSPERSTRCFRRSRFLAMGEPAYLLLLLLSLALGLVLAALGFVHHRDSPLYQ 600
 QY 601 ASGGPLACGGLVCLGLVCLSVLLFPQSPACLAQOPLSHPLPGCSTLFLQAAEIFV 660
 Db 601 ASGGPLACGGLVCLGLVCLSVLLFPQSPACLAQOPLSHPLPGCSTLFLQAAEIFV 660
 QY 661 ESEIPLSNADRLSGCLRGPMAMLVLLAMLVVALCTWLVAFPEPVYTDHMLPTEALV 720
 Db 661 ESEIPLSNADRLSGCLRGPMAMLVLLAMLVVALCTWLVAFPEPVYTDHMLPTEALV 720
 QY 721 HCTRSWVSFGIAHATNATLAFCLGFLVRSOPGCYNRARGLTFAMLAFTTWSEVP 780
 Db 721 HCTRSWVSFGIAHATNATLAFCLGFLVRSOPGCYNRARGLTFAMLAFTTWSEVP 780
 QY 781 LLANOVVLRPAVQMGALLCYGLTAAHFLPRCTLLMQPPLNPEFFLGSGPDAGQ 840
 Db 781 LLANOVVLRPAVQMGALLCYGLTAAHFLPRCTLLMQPPLNPEFFLGSGPDAGQ 840
 QY 841 NDGNTNGNCKHE 852
 Db 841 NDGNTNGNCKHE 852

RESULT 4
 AAU73184 standard; Protein; 852 AA.
 ID AAU73184 standard; Protein; 852 AA.
 AC AAU73184;
 DT 12-MAR-2002 (first entry)
 DE Human SAC1 polypeptide.
 DE Human SAC1 polypeptide.
 KM Human; mouse; SAC1; carbohydrate; sweetener; ethanol; alcoholism;
 KM obesity; diabetes; transgenic embryo; body tissue; body fluid; pancreas;
 KM blood; tongue; anorectic; antidiabetic; gene therapy;
 KM protein replacement therapy.
 OS Homo sapiens.
 PN W0200183749-A2.
 PD 08-NOV-2001.
 PF 25-APR-2001; 2001WO-US13387.
 PR 28-APR-2000; 2000US-200794P.
 PR 28-JUL-2000; 2000US-221419P.
 PR 10-NOV-2000; 2000US-247443P.
 PA (WARN) WARNER LAMBERT CO.
 PA (MONE-) MONELL CHEM SENSES CENT.
 PI Bachmanov AA, Beauchamp GK, Chatterjee A, De Jong PJ, Li S, Li X;
 PI Ohmen JD, Reed DR, Ross D, Tordoff MG;
 DR WPI; 2002-075162/10.
 DR N-Psdb; AAS97395.
 PT Novel isolated polypeptide comprising variant form of mouse or human
 PT SAC1 polypeptide, and is associated with altered preference for
 PT carbohydrates or other sweeteners, useful for preventing obesity,
 PT diabetes, alcoholism
 PS Claim 7; Page 44-45; 239pp; English.
 XX The invention relates to an isolated polypeptide, comprising a variant
 CC form of mouse or human SAC1 polypeptide. The variant form is associated
 CC with altered preference for carbohydrates, other sweeteners or ethanol.
 CC The polypeptide and its associated DNA sequence can be produced by
 CC recombinant techniques and is useful for preventing obesity, diabetes or
 CC alcoholism associated with SAC1 expression. The sequences are useful in
 CC screening for drugs and sweeteners. Recombinant cell lines and transgenic

CC embryos may be used in screening for and identifying agents that induce
 CC or repress function of SAC1. Predisposition to diabetes, obesity or
 CC alcoholism can be ascertained by testing any fluid or tissue of a human
 CC (such as blood, pancreas or tongue) for sequence variations of the SAC1
 CC gene. A sequence variation of the SAC1 locus may indicate a
 CC predisposition to diabetes, obesity and/or alcoholism and may provide a
 CC diagnostic mark. The polynucleotide can be detected in a biological
 CC sample by contacting the DNA with a probe to form a hybridization complex
 CC which is then detected. The sequences represent the human and mouse SAC1
 CC polypeptides.

Sequence 852 AA;

Query Match 99.7%; Score 4509; DB 23; Length 852;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 850; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGPAVLGLSLMALHPGAPCLCSOOLRMKGDVYLGLPFGAEAGLSRTRPSP 60
 Db 1 MLGPAVLGLSLMALHPGAPCLCSOOLRMKGDVYLGLPFGAEAGLSRTRPSP 60
 QY 61 VCTRESSNGLMALAMKAAVEEINKSDLLPGLRLGYDLFTDCSEPVYAMKPSLFLAKA 120
 Db 61 VCTRESSNGLMALAMKAAVEEINKSDLLPGLRLGYDLFTDCSEPVYAMKPSLFLAKA 120
 QY 121 GSRDIAACNTYQYQPRVLAIVGPHSSSELAATGKFFSEFLMPQVSYGASMLLSARETF 180
 Db 121 GSRDIAACNTYQYQPRVLAIVGPHSSSELAATGKFFSEFLMPQVSYGASMLLSARETF 180
 QY 181 PSFRTVSDRYQVLTAAAEILQEFGMNMYAALGSDDEXGROGLSTFSAALAGRCIAHE 240
 Db 181 PSFRTVSDRYQVLTAAAEILQEFGMNMYAALGSDDEXGROGLSTFSAALAGRCIAHE 240
 QY 241 GLVPLPRADDSRLGVQDYDLHOVNOSVQVYLLFASVHAHALENYSSISRLSPVYWAAS 300
 Db 241 GLVPLPRADDSRLGVQDYDLHOVNOSVQVYLLFASVHAHALENYSSISRLSPVYWAAS 300
 QY 301 EAMLTSDLYMGLPGMAQGTIVGFLORGAQLHEFPQYKTHLATDPAFCALGERDQ 360
 Db 301 EAMLTSDLYMGLPGMAQGTIVGFLORGAQLHEFPQYKTHLATDPAFCALGERDQ 360
 QY 361 LEEDVVGQRCPQDCITLONVSAGLNHHOTFSVYAAVSVAAQALNHTLQCNASGCPADP 420
 Db 361 LEEDVVGQRCPQDCITLONVSAGLNHHOTFSVYAAVSVAAQALNHTLQCNASGCPADP 420
 QY 421 VKPQDLLENMYLTFHVGGPLPRFDSSGNVDMETDKLWMOGSVPRLDHVGRENGSLRT 480
 Db 421 VKPQDLLENMYLTFHVGGPLPRFDSSGNVDMETDKLWMOGSVPRLDHVGRENGSLRT 480
 QY 481 ERLKTRMHTSDNQKPYSCRSRCCQGOYRVYGFHSCCYDCVDCAGSVRONPDIACTF 540
 Db 481 ERLKTRMHTSDNQKPYSCRSRCCQGOYRVYGFHSCCYDCVDCAGSVRONPDIACTF 540
 QY 541 CGODEMSPERSTRCFRRSRFLAMGEPAYLLLLLLSLALGLVLAALGFVHHRDSPLYQ 600
 Db 541 CGODEMSPERSTRCFRRSRFLAMGEPAYLLLLLLSLALGLVLAALGFVHHRDSPLYQ 600
 QY 601 ASGGPLACGGLVCLGLVCLSVLLFPQSPACLAQOPLSHPLPGCSTLFLQAAEIFV 660
 Db 601 ASGGPLACGGLVCLGLVCLSVLLFPQSPACLAQOPLSHPLPGCSTLFLQAAEIFV 660
 QY 661 ESEIPLSNADRLSGCLRGPMAMLVLLAMLVVALCTWLVAFPEPVYTDHMLPTEALV 720
 Db 661 ESEIPLSNADRLSGCLRGPMAMLVLLAMLVVALCTWLVAFPEPVYTDHMLPTEALV 720
 QY 721 HCTRSWVSFGIAHATNATLAFCLGFLVRSOPGCYNRARGLTFAMLAFTTWSEVP 780
 Db 721 HCTRSWVSFGIAHATNATLAFCLGFLVRSOPGCYNRARGLTFAMLAFTTWSEVP 780
 QY 781 LLANOVVLRPAVQMGALLCYGLTAAHFLPRCTLLMQPPLNPEFFLGSGPDAGQ 840
 Db 781 LLANOVVLRPAVQMGALLCYGLTAAHFLPRCTLLMQPPLNPEFFLGSGPDAGQ 840

QY 841 NDGNTGNKGHE 852
 |||||
 Db 841 NDGNTGNKGHE 852

RESULT 5
 AAEL0366
 ID AAEL0366 standard; Protein: 850 AA.
 AC AAEL0366;
 DT 10-DEC-2001 (first entry)
 DE Human taste-cell-specific G protein-coupled receptor, hTIR3 protein.
 XX
 XX Human taste-cell-specific G protein-coupled receptor; TIR3; drug;
 KM genetic modulation; pharmaceutical; taste sensation; food industry;
 KM sensory transduction; chromosome 1p36.2-1p36.33.
 OS Homo sapiens.
 XX
 XX W0200166563-A2.
 PD 13-SEP-2001.
 PF 07-MAR-2001; 2001MO-US07265.
 PR 07-MAR-2000; 2000US-0187546.
 PR 07-APR-2000; 2000US-0195536.
 PR 06-JUN-2000; 2000US-0209840.
 PR 23-JUN-2000; 2000US-0214213.
 PR 17-AUG-2000; 2000US-0226448.
 PR 03-JAN-2001; 2001US-0259227.
 XX
 XX (SEMO-) SENOMYX INC.
 PA
 PI Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
 DR WPI: 2001-582267/65.
 DR N-PSDB: AAD17509, AAD17517.
 XX
 XX New mammalian taste-cell-specific G protein-coupled receptor
 PT polypeptides for identifying compounds that modulate taste signaling
 PT are useful in food, to modulate the sweet taste of foods or drugs -
 XX
 XX Claim 152; Page 73-74; 119pp; English.

CC The invention relates to mammalian taste-cell-specific G protein-coupled
 CC receptors, TIR and their corresponding cDNA molecules. Taste receptors,
 CC TIR are useful for screening compounds which are used to activate or
 CC modulate chemosensory transduction, such as taste sensation. The
 CC identification and isolation of novel taste receptors and taste
 CC signalling molecules allow for new methods of chemical and genetic
 CC modulation of taste transduction pathways. The taste modulating
 CC compounds are useful in pharmaceuticals and food industries to improve
 CC the taste of a variety of consumer products, or to block undesirable
 CC tastes, e.g., in certain pharmaceuticals. TIRs are also useful in
 CC biochemical assay for identifying tastant (TIR) ligands having binding
 CC specificity for TIR involved in taste signalling. The present sequence is
 CC human taste-cell-specific G protein-coupled receptor, hTIR3 protein.
 CC Human TIR3 gene is localised on chromosome 1p36.2-1p36.33.
 XX
 XX Sequence 850 AA;

Query Match 99.5%; Score 4500; DB 22; Length 850;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 84; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 MLCGPAVLGSLMALHPGTGAPLCTSOOLRMKGDYVIGLPLGEAEPAGRSRTRPSSP 60
 |||||
 Db 1 MLCGPAVLGSLMALHPGTGAPLCTSOOLRMKGDYVIGLPLGEAEPAGRSRTRPSSP 60

QY 61 VCFRFSNGLLMALAMKAAVEINNKSDLLPGLRLGYDLFTDCEPVPVAKPSIMFLAKA 120

Db 61 VCFRFSNGLLMALAMKAAVEINNKSDLLPGLRLGYDLFTDCEPVPVAKPSIMFLAKA 120
 |||||
 QY 121 GSRDIAVCNTYOYOPRYLVAVIGHSSELAMVYGFSEFLMPOVSYGASMEILSARETF 180
 |||||
 Db 121 GSRDIAVCNTYOYOPRYLVAVIGHSSELAMVYGFSEFLMPOVSYGASMEILSARETF 178
 |||||
 QY 181 PSFRTVPSPDRVOLTAAEELLOEFGNMVAALSGDDEYGRQGLSIFSAALAAAGICIAHE 240
 |||||
 Db 179 PSFRTVPSPDRVOLTAAEELLOEFGNMVAALSGDDEYGRQGLSIFSAALAAAGICIAHE 238
 |||||
 QY 241 GLVPLPRADSRIGKVDVYHOVNOSSVQVYLLFASVHAHALFNISISRLSPKVVAS 300
 |||||
 Db 239 GLVPLPRADSRIGKVDVYHOVNOSSVQVYLLFASVHAHALFNISISRLSPKVVAS 298
 |||||
 QY 301 EAWLTSLVWGLPMAOMGTVLGFLQGAOLHEPPOVYKTHLALADPAPCSALGEREOG 360
 |||||
 Db 299 EAWLTSLVWGLPMAOMGTVLGFLQGAOLHEPPOVYKTHLALADPAPCSALGEREOG 358
 |||||
 QY 361 LEEDVVGQRPCDCITLQNVASAGLNHQTFSYAAVYVAQALHNTLQCNASGCPAODP 420
 |||||
 Db 359 LEEDVVGQRPCDCITLQNVASAGLNHQTFSYAAVYVAQALHNTLQCNASGCPAODP 418
 |||||
 QY 421 VKPQQLLENNYNTLFHYGILPLRFDSSGVNDMEYDLKIMWQGSVPLHYGRNGSLRT 480
 |||||
 Db 419 VKPQQLLENNYNTLFHYGILPLRFDSSGVNDMEYDLKIMWQGSVPLHYGRNGSLRT 478
 |||||
 QY 481 ERLKIRHMTSDNOKPVSRGSCNOCEGOVRRYKGFHSQCYOCDEAGSYRONPDIACTE 540
 |||||
 Db 479 ERLKIRHMTSDNOKPVSRGSCNOCEGOVRRYKGFHSQCYOCDEAGSYRONPDIACTE 538
 |||||
 QY 541 CGODEMSPERSTRCFRRRSFELANGEPVALLLLLSIALGLVLAALGLFVHHRDSEPLVQ 600
 |||||
 Db 539 CGODEMSPERSTRCFRRRSFELANGEPVALLLLLSIALGLVLAALGLFVHHRDSEPLVQ 598
 |||||
 QY 601 ASGGPLACFGILVCLGLVCLSVLLFPGQSPARCLAQOPLSHLPLTGLSTLFLQAAEIV 660
 |||||
 Db 599 ASGGPLACFGILVCLGLVCLSVLLFPGQSPARCLAQOPLSHLPLTGLSTLFLQAAEIV 658
 |||||
 QY 661 ESELPLSMADRLSGCLGRPNAMLVLLAMLEVALCWYVYAPPEVVTOMHMLPTFALV 720
 |||||
 Db 659 ESELPLSMADRLSGCLGRPNAMLVLLAMLEVALCWYVYAPPEVVTOMHMLPTFALV 718
 |||||
 QY 721 HCRTRSWSPGLAHATNATLAFCLFGLTFLVRSOPGCYNRARGITFAMLAYFITWSEVP 780
 |||||
 Db 719 HCRTRSWSPGLAHATNATLAFCLFGLTFLVRSOPGCYNRARGITFAMLAYFITWSEVP 778
 |||||
 QY 781 LLANVQVVLPRAYOMGALLLCVGLTILAAFLHPRCYLLMRPGINTPEFFLGGGDAQOQ 840
 |||||
 Db 779 LLANVQVVLPRAYOMGALLLCVGLTILAAFLHPRCYLLMRPGINTPEFFLGGGDAQOQ 838
 |||||
 QY 841 NDGNTGNKGHE 852
 |||||
 Db 839 NDGNTGNKGHE 850

RESULT 6
 ABB77318
 ID ABB77318 standard; Protein: 863 AA.
 AC ABB77318;
 DT 17-JUN-2002 (first entry)
 DE Human G-protein coupled receptor SRO ID NO 2.
 XX
 XX
 XX Human; GPCR; G-protein coupled receptor; receptor; anti-HIV; antitumour;
 KM antiinflammatory; antiallergic; antianemic; antiasthmatic; virucide;
 KM immunosuppressive; dermatological; nephrotropic; antigout; antithyroid;
 KM cytostatic; neuroprotective; osteopathic; antipsoriatic; antineumatic;
 KM antiarthritic; thyromimetic; antiulcer; ophthalmological; antibacterial;
 KM fungicide; antiparasitic; protozoacide; antihelmintic; antidiabetic;
 KM antiarteriosclerotic; hepatotropic; anticonvulsant; anorectic; metabolic;

antimetetic; antidarrhoeic; neuroleptic; cerebroprotective; nootropic; antiparkinsonian; depilatory; tranquilizer; hypotensive; vasotropic; cardiatic; antiangiial; vulnerary; proliferative disorder; cancer; neurological disorder; Alzheimer's disease; Huntington's disease; Parkinson's disease; multiple sclerosis; meningitis; prion; cardiovascular disorder; acquired immunodeficiency syndrome; AIDS; Crohn's disease; diabetes mellitus; rheumatoid arthritis.

XX
OS Homo sapiens.

XX
PN WO200198323-A2.

XX
PD 27-DEC-2001.

XX
PF 15-JUN-2001; 2001WO-US91354.

XX
PR 16-JUN-2000; 2000US-212483P.

XX
PR 23-JUN-2000; 2000US-213950P.

PR 26-JUN-2000; 2000US-214062P.

PR 07-JUL-2000; 2000US-216555P.

PR 14-JUL-2000; 2000US-218936P.

XX
PR 19-JUL-2000; 2000US-219154P.

XX
PA (IMCY-) INCYTE GENOMICS INC.

XX
XX Lal P, Graul R, Hafalia AA, Walla NK, Thornton M, Nguyen DB;
P1 Lu Y, Gandhi AR, Patterson C, Kallilick DA, Baughn MR, Ramkumar J;
P1 Tribouley CM, Lee EA, Ding L, Burford N, Yao MG, Yang J;
P1 Griffin JA;

XX
DR WPI: 2002-139780/18.

XX
DR N-PSDB; ABL55952.

XX
PT Novel G-protein coupled receptor protein and polynucleotides useful for
PT diagnosing, treating or preventing disorders of cell proliferation e.g.
PT cancer, neurological and genetic disorder e.g. thalassemia

XX
PS Claim 47; Page 105-107; 121pp; English.

XX
XX The invention relates to a G-protein coupled receptor protein (GCREC)
CC polypeptide (AB877317-AB877326). The GCREC is useful for screening an
CC agonist/antagonist of GCREC, a compound that specifically binds to GCREC
CC or that modulates the activity of GCREC. GCREC is also useful as an
CC immunogen for preparing antibodies which are useful for diagnosing a
CC condition of disease associated with expression of GCREC in a subject,
CC for detecting and purifying GCREC from a sample. The GCREC encoding
CC polynucleotide (ABL55951-ABL55960) is useful for screening for a compound
CC effective in altering expression of GCREC. GCREC is also useful for
CC assessing toxicity of a test compound useful for treating a disease or
CC condition associated with decreased expression or overexpression of
CC functional GCREC. Examples of disorders include cell proliferative
CC disorder such as arteriosclerosis, atherosclerosis, hepatitis, mixed
CC connective tissue disease (MCTD), psoriasis and cancer including
CC adenocarcinoma, leukaemia; a neurological disorder such as epilepsy,
CC stroke, Alzheimer's disease, Huntington's disease, dementia, Parkinson's
CC disease, retinitis pigmentosa, multiple sclerosis, bacterial and viral
CC meningitis, abscess, subdural empyema; prion disease including kuru,
CC Creutzfeldt-Jakob disease; fatal familial insomnia, neurofibromatosis,
CC tuberous sclerosis, cerebral palsy, polymyositis; inherited, metabolic,
CC endocrine, and toxic myopathies; myasthenia gravis, periodic paralyses;
CC mental disorders including mood, anxiety, and schizophrenic disorders;
CC seasonal affective disorder (SAD); akathisia, amnesia, cataplexia,
CC diabetic neuropathy, tardive dyskinesia, dystonia, paranoid psychoses,
CC Tourette's disorder; cardiovascular disorders such as hypertension,
CC vasculitis, vascular tumours, congestive heart failure, ischaemic heart
CC disease, myocardial infarction, calcific aortic valve stenosis,
CC infective endocarditis, endocarditis of systemic lupus erythematosus,
CC cardiac transplantation; gastrointestinal disorder such as dysphagia,
CC gastritis, anorexia, nausea, emesis, abdominal angina, infections of
CC the intestinal tract, peptic ulcer, hepatitis, cirrhosis, diarrhoea,
CC acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, Reye's
CC syndrome, liver infarction, an autoimmune/inflammatory disorder such as
CC AIDS, Addison's disease, adult respiratory distress syndrome,

CC	allergies, amyloidosis, anaemia, asthma, atherosclerosis, autoimmune
CC	thyroiditis, bronchitis, Crohn's disease, diabetes mellitus, thyroiditis,
CC	Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis,
CC	multiple sclerosis, myasthenia gravis, myocardial or pericardial
CC	inflammation, osteoporosis, psoriasis, rheumatoid arthritis,
CC	scleroderma, Sjogren's syndrome, systemic lupus erythematosus, uveitis,
CC	viral, bacterial, fungal, parasitic, protozoal, and hematologic
CC	infections, and trauma; a metabolic disorders such as diabetes,
CC	obesity, and osteoporosis. The GCPEC encoding polynucleotide is useful
CC	for gene therapy and for creating knock in humanised animals (pigs) or
CC	transgenic animals (mice or rats) to model human disease.
XX	
50	Sequence 863 AA;
Query Match	99.4%; Score 4496.5; DB 23; Length 863;
Best Local Similarity	98.6%; Pred. No. 0;
Matches 851;	Conservative 0; Mismatches 1; Indels 11; Gaps 1
OY	1 MGPVAVGISIMWALLHPGTGAPLCUSQQLRMKGDDYVGLFPLGEAEAGLSRRPSSP 60
DB	1 MGPVAVGISIMWALLHPGTGAPLCUSQQLRMKGDDYVGLFPLGEAEAGLSRRPSSP 60
OY	61 VCTRRSSNGLMALKMAVEEINNKSDLLPGLRLGYDLPDCCSEYVAMKRSIMFLA 120
DB	61 VCTRRSSNGLMALKMAVEEINNKSDLLPGLRLGYDLPDCCSEYVAMKRSIMFLA 120
OY	121 GSRDLAAYCNTYOQPRVLAIVGPHSSSELAATYGRKFFSEFLMPQVSYGASMETLSARETF 180
DB	121 GSRDLAAYCNTYOQPRVLAIVGPHSSSELAATYGRKFFSEFLMPQVSYGASMETLSARETF 180
OY	181 PSFEFTVPSDRVOLTAAELLQDFGNWVAALGSDDEYRGQGISFSAALAAKGICIAHE 240
DB	181 PSFEFTVPSDRVOLTAAELLQDFGNWVAALGSDDEYRGQGISFSAALAAKGICIAHE 240
OY	241 GLVPLPRADDSRLGKQVDLVHVNQSSVQVLLFASVHAHLNFNSISRLSPKTVAS 300
DB	241 GLVPLPRADDSRLGKQVDLVHVNQSSVQVLLFASVHAHLNFNSISRLSPKTVAS 300
OY	301 EAMLTSDIYMGIPGMAOMGTVLGFTQRCAGQLHEFPQVYTHALATDPAFCALGEREG 360
DB	301 EAMLTSDIYMGIPGMAOMGTVLGFTQRCAGQLHEFPQVYTHALATDPAFCALGEREG 360
OY	361 LEEEDVVGRCPOCCDCTTQNVASAGLNHQTFSVYAAVSAOALHTLQCNASGCPADP 420
DB	361 LEEEDVVGRCPOCCDCTTQNVASAGLNHQTFSVYAAVSAOALHTLQCNASGCPADP 420
OY	421 VEPQQLLEMYNLTFHYGGLPLRFSSGQVMDMEYDLKLVNQGSSVRLMDVGRNSLT 480
DB	421 VEPQQLLEMYNLTFHYGGLPLRFSSGQVMDMEYDLKLVNQGSSVRLMDVGRNSLT 480
OY	481 ERLKTRMHTSDN-----OKFVSRCSRCOCQGOVRYKGFPSCCYDCVDCAGSY 529
DB	481 ERLKTRMHTSDN-----OKFVSRCSRCOCQGOVRYKGFPSCCYDCVDCAGSY 540
OY	530 RQNPDDIACFCGODEWSPDERSTRCFRRSRLANGEPVLLLLLLSLALGIVLAALG 589
DB	541 RQNPDDIACFCGODEWSPDERSTRCFRRSRLANGEPVLLLLLLSLALGIVLAALG 600
OY	590 FVHHNDSPLVOASGGPLACFGLVCLGLVLSVLLPFGQPSPARCLAAQPLSHLPLTGCS 649
DB	601 FVHHNDSPLVOASGGPLACFGLVCLGLVLSVLLPFGQPSPARCLAAQPLSHLPLTGCS 660
OY	650 TFLQAAELFEVSESLPLSMADRLSGLRPMAMIVLLMLVEVALCTYVLAEPVEYT 709
DB	661 TFLQAAELFEVSESLPLSMADRLSGLRPMAMIVLLMLVEVALCTYVLAEPVEYT 720
OY	710 DMHMLPTEALVACRTRSWVSFGLAHATNTATLFLCFGLFVRSQPGCYNRAAGLIFAML 769
DB	721 DMHMLPTEALVACRTRSWVSFGLAHATNTATLFLCFGLFVRSQPGCYNRAAGLIFAML 780
OY	770 AVEITWVSFVPLLANVQVLRPAVOMGALLLCVGLIAAFLPRCYLLMRQGLNTPPEF 829
DB	781 AVEITWVSFVPLLANVQVLRPAVOMGALLLCVGLIAAFLPRCYLLMRQGLNTPPEF 840

QY 830 IGGGPGDAGGNDGNTGNCKHE 852
 ||||||||||||||||||
 DB 841 IGGGPGDAGGNDGNTGNCKHE 863

RESULT 7
 AAU80493
 ID AAU80493 standard; Protein; 855 AA.
 XX

AC AAU80493;

DT 12-MAR-2002 (first entry)

DE Human G-coupled receptor (GCREC) protein, Seq ID No 1.

XX Human: cytosolic; neuroprotective; immunosuppressant; nocitropic;
 KM anti-inflammatory; anti-viral; gastrointestinal; cardiovascular;
 KM cerebroprotective; G-coupled receptor; cell proliferative disease;
 KM lymphoma; leukemia; breast cancer; cirrhosis; neurological disorder;
 KM stroke; Alzheimer's disease; multiple sclerosis; mental retardation;
 KM cardiovascular disease; atherosclerosis; angina pectoris; indigestion;
 KM congestive heart failure; gastrointestinal disorder; dysphagia; AIDS;
 KM gastritis; autoimmune disorder; inflammatory disorder; Crohn's disease;
 KM systemic lupus erythematosus; metabolic disorder; diabetes; obesity;
 KM viral infection; herpesvirus; parvovirus;
 KM acquired immune deficiency syndrome.

OS Homo sapiens.

PN WO200190359-A2.

PD 29-NOV-2001.

PE 22-MAY-2001; 2001MO-US16833.

PR 22-MAY-2000; 2000US-206222P.

PR 23-MAY-2000; 2000US-207476P.

PR 02-JUN-2000; 2000US-208834P.

PR 07-JUN-2000; 2000US-208861P.

PR 07-JUN-2000; 2000US-209868P.

XX (INCY-) INCYTE GENOMICS INC.

PI Patterson C, Tribouley CM, Yao MG, Griffin JA, Thornton M, Lu Y;

PI Kallick DA, Gandhi AR, Au-Young J;

XX WPI: 2002-106199/14.

DR N-PSDB; ABR16615.

XX New G-protein coupled receptors useful for treating or preventing cell

PT proliferative (e.g. leukemia), neurological (e.g. stroke),

PT cardiovascular or autoimmune/inflammatory disorders -

XX Claim 1; Page 116-118; 148pp; English.

PS The invention relates to a novel human G-coupled receptor (I). (I) and

XX its corresponding polynucleotides are useful for diagnosing, treating or

CC preventing cell proliferative diseases (e.g. lymphoma, leukemia, breast

CC cancer or cirrhosis), neurological disorders (e.g. stroke, Alzheimer's

CC disease, multiple sclerosis or mental retardation), cardiovascular

CC diseases (e.g. atherosclerosis, angina pectoris or congestive heart

CC failure), gastrointestinal disorders (e.g. dysphagia, indigestion or

CC gastritis), autoimmune/inflammatory disorders (e.g. AIDS, Crohn's disease

CC or systemic lupus erythematosus) or metabolic disorders (e.g. diabetes

CC or obesity), or viral infections (e.g. infection by herpesvirus or

CC parvovirus). AAU80493-AAU80515 represent novel human G-coupled

CC receptor amino acid sequences of the invention.

XX Sequence 855 AA;

SQ Query Match 82.0%; Score 3708; DB 23; Length 855;

Best Local Similarity 81.4%; Pred. No. 0;

Matches 738; Conservative 5; Mismatches 38; Indels 126; Gaps 10;

QY 1 MLGPAYGLSLMALHPTGAPLCLSOQLRKKGQYVGLGFLPGAEAEAGLRSTRSSP 60

DB 1 MLGPAYGLSLMALHPTGAPLCLSOQLRKKGQYVGLGFLPGAEAEAGLRSTRSSP 60

QY 61 VCTRSSNGLLMALAMKAAVEEINKSDLLPGLRGYDLFDYCEPVPYANKPSIMFLAKA 120

DB 61 VCTRSSNGLLMALAMKAAVEEINKSDLLPGLRGYDLFDYCEPVPYANKPSIMFLAKA 120

QY 121 GSRDIAVCNTYQYQPVYLAIVGPHSSELMAYGKFFSFLMP----- 163

DB 121 GSRDIAVCNTYQYQPVYLAIVGPHSSELMAYGKFFSFLMP----- 163

QY 164 -----QVSYGSMELLSARETFPSFFRTVPDRVQLTAALLOEFGNNVAAL 212

DB 164 -----QVSYGSMELLSARETFPSFFRTVPDRVQLTAALLOEFGNNVAAL 212

QY 181 APVSGDASWPLQSYSGSMELLSARETFPSFFRTVPDRVQLTAALLOEFGNNVAAL 240

DB 181 APVSGDASWPLQSYSGSMELLSARETFPSFFRTVPDRVQLTAALLOEFGNNVAAL 240

QY 213 GSDDYGRGGLSIFSAALAAAGICIAHEGLVPLPRADDSRLGRVQDVLHOVNOSSVQVTL 272

DB 213 GSDDYGRGGLSIFSAALAAAGICIAHEGLVPLPRADDSRLGRVQDVLHOVNOSSVQVTL 272

QY 241 GSDDYGRGGLSIFSAALAAAGICIAHEGLVPLPRADDSRLGRVQDVLHOVNOSSVQVTL 291

DB 241 GSDDYGRGGLSIFSAALAAAGICIAHEGLVPLPRADDSRLGRVQDVLHOVNOSSVQVTL 291

QY 273 LFASVHAHALFNYSISRSPKRVVASEAMVTSIDVMSLPGMONGTVLGLFQRGALH 332

DB 273 LFASVHAHALFNYSISRSPKRVVASEAMVTSIDVMSLPGMONGTVLGLFQRGALH 332

QY 292 LFASVHAHALFNYSISRSPKRVVASEAMVTSIDVMSLPGMONGTVLGLFQRGALH 351

DB 292 LFASVHAHALFNYSISRSPKRVVASEAMVTSIDVMSLPGMONGTVLGLFQRGALH 351

QY 333 EFPQYVKTALATDPAPFCALGEREGLEDVVGRCPODCITLONVSAGLNHHOTFS 392

DB 333 EFPQYVKTALATDPAPFCALGEREGLEDVVGRCPODCITLONVSAGLNHHOTFS 392

QY 352 EFPQYVKTALATDPAPFCALGEREGLEDVVGRCPODCITLONVSAGLNHHOTFS 399

DB 352 EFPQYVKTALATDPAPFCALGEREGLEDVVGRCPODCITLONVSAGLNHHOTFS 399

QY 393 VYAAVYSVAOLNNTLOCNASGCPADPVKPMOLLENNYLTFFVGGVPLRFDSGGVNDM 452

DB 393 VYAAVYSVAOLNNTLOCNASGCPADPVKPMOLLENNYLTFFVGGVPLRFDSGGVNDM 452

QY 400 -----RAQLNHTLLOCNASGCPADPVKPMOLLENNYLTFFVGGVPLRFDSGGVNDM 452

DB 400 -----RAQLNHTLLOCNASGCPADPVKPMOLLENNYLTFFVGGVPLRFDSGGVNDM 452

QY 453 EYDLKLMWOGSVPRLLDVGFRNGSLTERLTKIMWHSN-----QKVSRCSSR 501

DB 453 EYDLKLMWOGSVPRLLDVGFRNGSLTERLTKIMWHSN-----QKVSRCSSR 501

QY 453 EYDLKLMWOGSVPRLLDVGFRNGSLTERLTKIMWHSN-----QKVSRCSSR 512

DB 453 EYDLKLMWOGSVPRLLDVGFRNGSLTERLTKIMWHSN-----QKVSRCSSR 512

QY 502 QCGQGVRRVKGFSHCYDCVDCBAGSYRONPDIACTFCGDEWSPSRSTRCFRRSRF 561

DB 502 QCGQGVRRVKGFSHCYDCVDCBAGSYRONPDIACTFCGDEWSPSRSTRCFRRSRF 561

QY 513 QCGQGVRRVKGFSHCYDCVDCBAGSYRONPDIACTFCGDEWSPSRSTRCFRRSRF 572

DB 513 QCGQGVRRVKGFSHCYDCVDCBAGSYRONPDIACTFCGDEWSPSRSTRCFRRSRF 572

QY 562 LAMGEPAVLLLLLSLALGLVLAALGFVHHRSPLVQASGGGLACGLVGLVCLSV 621

DB 562 LAMGEPAVLLLLLSLALGLVLAALGFVHHRSPLVQASGGGLACGLVGLVCLSV 621

QY 573 LAMGEPAVLLLLLSLALGLVLAALGFVHHRSPLVQASGGGLACGLVGLVCLSV 632

DB 573 LAMGEPAVLLLLLSLALGLVLAALGFVHHRSPLVQASGGGLACGLVGLVCLSV 632

QY 622 LFPGQSPARCLAQOPLSHLPLTGCLSTFLQAAEFVSESLPLSWADRLSGCLRGPWA 681

DB 622 LFPGQSPARCLAQOPLSHLPLTGCLSTFLQAAEFVSESLPLSWADRLSGCLRGPWA 681

QY 633 LFPGQSPARCLAQOPLSHLPLTGCLSTFLQAAEFVSESLPLSWADRLSGCLRGPWA 692

DB 633 LFPGQSPARCLAQOPLSHLPLTGCLSTFLQAAEFVSESLPLSWADRLSGCLRGPWA 692

QY 682 WLVLVLAALVEVALCTWYLVAFPEVVTDMHMLPTEALVHCRTRSWVSGLAHAT----- 736

DB 682 WLVLVLAALVEVALCTWYLVAFPEVVTDMHMLPTEALVHCRTRSWVSGLAHAT----- 736

QY 693 WLVLVLAALVEVALCTWYLVAFPEVVT-----GLAHAAGGAG 731

DB 693 WLVLVLAALVEVALCTWYLVAFPEVVT-----GLAHAAGGAG 731

QY 737 -----NATLAFLCIGTFLVRSQPC-----CYNRAGLTFAMLAY 771

DB 737 -----NATLAFLCIGTFLVRSQPC-----CYNRAGLTFAMLAY 771

QY 732 ALPHLLLGQLRPSARHATTAFLCTGFH-----PGAEPAGLPQPCVASHICHAGLLHH 786

DB 732 ALPHLLLGQLRPSARHATTAFLCTGFH-----PGAEPAGLPQPCVASHICHAGLLHH 786

QY 772 FITWSEFVPLIAN-VQVVLPPAVQMGALLLCVGLIILAFLHPLRCYLMLRPGLTPEFFL 830

DB 772 FITWSEFVPLIAN-VQVVLPPAVQMGALLLCVGLIILAFLHPLRCYLMLRPGLTPEFFL 830

QY 787 --TSHHVPPLLAQAGSHSGPAVQMGALLLCVGLIILAFLHPLRCYLMLRPGLTPEFFL 844

DB 787 --TSHHVPPLLAQAGSHSGPAVQMGALLLCVGLIILAFLHPLRCYLMLRPGLTPEFFL 844

QY 831 GGGPGDA 837

DB 831 GGGPGDA 851

RESULT 8

AAE10371

ID AAE10371 standard; Protein; 858 AA.

XX

AC AAE10371.

XX

DT 10-DEC-2001 (first entry)
 XX Rat taste receptor, t1R3 protein.
 XX Rat, taste-cell-specific G protein-coupled receptor; t1R3; drug;
 KM genetic modulation; pharmaceutical; taste sensation; food industry;
 KM Chemosensory transduction.
 XX
 OS Rattus sp.
 XX
 PN W020016563-A2.
 PD 13-SEP-2001.
 PF 07-MAR-2001; 2001WO-US07265.
 PR 07-MAR-2000; 2000US-0187546.
 PR 07-APR-2000; 2000US-0195536.
 PR 06-JUN-2000; 2000US-0209840.
 PR 23-JUN-2000; 2000US-0214213.
 PR 17-AUG-2000; 2000US-0226448.
 PR 03-JAN-2001; 2001US-0259227.
 XX
 PA (SEMO-) SENOMTX INC.
 PI Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
 DR N-PSDB; AAD17514.
 XX
 PT New mammalian taste-cell-specific G protein-coupled receptor
 PT polypeptides for identifying compounds that modulate taste signaling
 PT are useful in food, to modulate the sweet taste of foods or drugs -
 PS Claim 152; Page 77; 119pp; English.
 XX
 CC The invention relates to mammalian taste-cell-specific G protein-coupled
 CC receptors, t1R and their corresponding cDNA molecules. Taste receptors,
 CC t1R are useful for screening compounds which are used to activate or
 CC modulate chemosensory transduction, such as taste sensation. The
 CC identification and isolation of novel taste receptors and taste
 CC signaling molecules allow for new methods of chemical and genetic
 CC modulation of taste transduction pathways. The taste modulating
 CC compounds are useful in pharmaceuticals and food industries to improve
 CC the taste of a variety of consumer products, or to block undesirable
 CC tastes, e.g., in certain pharmaceuticals. t1R's are also useful in
 CC biochemical assay for identifying tastant (t1R) ligands having binding
 CC specificity for t1R involved in taste signaling. The present sequence is
 CC rat taste-cell-specific G protein-coupled receptor, t1R3 protein.
 XX
 SQ Sequence 858 AA:
 Query Match 71.9%; Score 3255; DB 22; Length 858;
 Best Local Similarity 72.9%; Pred. No. 2.6e-304;
 Matches 623; Conservative 74; Mismatches 144; Indels 14; Gaps 4;
 QY 1 MGPVAVIGSLMALHPGTGAPLCLSQOLRMKGDVYIGLPLGEAEAGLRPRPSP 60
 DB 1 MGLAVIIGSLAFLGELGSSSLCSQPKAGDVIIGLFLGTEETLWNRQPKNI 60
 QY 61 VCTRSSNGLLMALAKMAVEEINKSDLLPGILRLGYDLFDTCSEPVAMKPSLFLAKA 120
 DB 61 LCTRESPLGLFLAMAKMAVEEINNGSALLPGILRLGYDLFDTCSEPVAMKPSLFLAKV 120
 QY 121 GSRDIAAYCNYOYOPRVAVIGPSSSELAMTYGKFFSFLMPQVSYGSMELSLARETF 180
 DB 121 GSQSTIAAYCNYOYOPRVAVIGPSSSELALITGKFFSFLMPQVSYGSMELSLARETF 180
 QY 181 PSFFFTVPSPDRYQVLAALAELEDFGNNVVAALGSDDEYRGOSLTFSAALAAAGCICIAHE 240
 DB 181 PSFFFTVPSPDRYQVLAALAELEDFGNNVVAALGSDDEYRGOSLTFSAALAAAGCICIAHE 240
 QY 241 GLVPLPRADSRLLGKVVQVLYHVNQSSVQVLLLFASVHAHALFNYSISSRLSPKVVAS 300

DB 241 GLVPLPRADSRLLGKVVQVLYHVNQSSVQVLLLFASVHAHALFNYSISSRLSPKVVAS 300
 QY 301 EAMLISDLYMGPLGMAQMTVIGFLORGAQLHEFPQYKTHIALATDPAFCSALGEREG 360
 DB 301 ESMLSLWLTPLPNTARVGTIVGFLORGAQLHEFPQYKTHIALATDPAFCSALGEREG 360
 QY 361 LEEDVVGORPCQDCITLQNVASGLN-----HHQFSVYAAYSVQAALHNTLQCA 412
 DB 360 LEEVWGPSCQDYLMLQNLSSGLMONTISAGQLHQITATTAAYSVQAALHNTLQCA 419
 QY 413 SCGPADPYKPMQLLENMNTLTFHVGLPLRFDDSGNVDMEDLKLWVQGSVPRLADYV 472
 DB 420 SHCHTSEPVPQMLENNMNSFRARDLTLQDAGSVDMEDLKLWVQGSVPRLADYV 479
 QY 473 RENGSLRTERLKRIMWTSNOKPVSRCSQCOBGVRYKGFHSCCYDCVCEAGSYRON 532
 DB 480 FENGTLQLOHSMKYM--PENQVPVSCSHOCKDQGVRRYKGFHSCCYDCVCEAGSYRON 537
 QY 533 PDICTFCGODEMSPERSTRCFRRSRPLANGEPVALLLLSLALGLVLAALGLFVH 592
 DB 538 PDDFTCTPCGKDOMSPEKSTTCLPRPKFLANGEPVALLLLSLALGLVLAALGLFVH 597
 QY 593 HRDSPLVQASGGPLACFGLVCLGLVCLSVLFPGGQSPARCIAQOPLSLPTGCLSTLF 652
 DB 598 YWDSPLVQASGGSLFCFGLICLGLFCLSVLFPGRFRSASCLAQPMALPLTGLCLSTLF 657
 QY 653 LQAAELFVSELPPLSWADLSGCLRPMALVYLLAMLEVLAALCTVYLVAFPEVYTDH 712
 DB 658 LQAAELFVSELPPLSWANLCSYLRPMALVYLLATLEAALCAVYLMAFPEVYTDH 717
 QY 713 MLPTALVCRFRSWSFGLAHATNATLAFCLFGLFELVRSQPCYNNRARGLFAMLA 772
 DB 718 VLPTVLEHCRMRSWSLGVIHITNVLAFCLFGLFELVRSQPCYNNRARGLFAMLA 777
 QY 773 TIWVSFVPLLANVQVAVQMGALLLCVIGLILAFHLPRCYLIMRQGLMTPPEFLIG 832
 DB 778 TIWVSFVPLLANVQVAVQMGALLLCVIGLILAFHLPRCYLIMRQGLMTPPEFLIG 837
 QY 833 GEGDAQOGDNGTGN 847
 DB 838 SPKEA---SDGNSGS 849
 RESULT 9
 AAU73183
 ID AAU73183 standard; Protein: 858 AA.
 AC AAU73183;
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DE Murine SACL polypeptide.
 XX
 KM Human; mouse; SACL; carbohydrate; sweetener; ethanol; alcoholism;
 KM obesity; diabetes; transgenic embryo; body tissue; body fluid; pancreas;
 KM blood; tongue; anorectic; antidiabetic; gene therapy;
 KM protein replacement therapy.
 OS Mus sp.
 XX
 PN W0200183749-A2.
 PD 08-NOV-2001.
 PF 25-APR-2001; 2001WO-US13387.
 PR 28-APR-2000; 2000US-2007949.
 PR 28-JUL-2000; 2000US-221419P.
 PR 10-NOV-2000; 2000US-247443P.
 XX
 PA (WARN) WARNER LAMBERT CO.
 PA (MONE-) MONELL CHEM SENSES CENT.

OY 818 MROGLNTEPEF 829
 DB 1127 LCRPELNTEHF 1138

RESULT 13
 ID AAY45028 standard; Protein: 840 AA.

AC AAY45028;
 DT 31-MAY-2000 (first entry)

DE Rat sensory transduction G-protein coupled receptor-B3 variant #3.
 XX
 XX Rat; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
 KW sensory cell; taste receptor cell; screen; taste modulator;
 KW pharmaceutical; food industry; taste topographic map; tongue; variant.
 XX
 OS Rattus sp.
 OS Synthetic.

XX
 FH Key Location/Qualifiers
 FT Misc-difference 90 /note="Wild type Ala substituted by Gly"
 XX
 PN W0200006592-A1.
 XX
 PD 10-FEB-2000.
 XX
 PF 27-JUL-1999; 99WO-US17099.
 XX
 PR 28-JUL-1998; 98US-0094465.
 XX
 PA (REGC) UNIV CALIFORNIA.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Zuker CS, Adler JE, Lindemeyer J, Ryba N, Hoon M;
 XX WPT: 2000-205451/18.
 DR
 XX
 PT New isolated sensory transduction G-protein coupled receptor, useful
 PT for developing products for use in studying and modulating the taste
 PT transduction pathway -
 XX
 PS Disclosure: Page -: 83pp; English.

XX The present sequence is a polymorphic variant #3 of rat taste cell
 CC specific G-protein coupled receptor, GPCR-B3 which is involved in
 CC sensory transduction. GPCR-B3 is specifically expressed in foliate
 CC and fungiform cells, with lower expression in circumvallate taste
 CC receptor cells of the tongue. The protein has an extracellular domain,
 CC seven transmembrane domains and an intracellular domain.
 CC The GPCR-B3 sequence is used to screen compounds that modulate sensory
 CC signalling in taste cells, especially taste modulators useful in
 CC pharmaceutical and food industries to customise taste. The sequence
 CC can also be used as probe for identifying taste cells and
 CC subsets of taste receptor cells such as foliate, fungiform and
 CC circumvallate. Such probes are also useful to generate taste
 CC topographic maps that elucidate the relationship between the taste
 CC cells of the tongue and sensory neurons leading to taste centres
 CC in the brain.
 CC Note: The present sequence is not given in the specification but is
 CC derived from rat GPCR-B3 sequence shown in page 75 (AAY45021).
 XX
 SQ Sequence 840 AA;

Query Match: 26.7%; Score 1209.5; DB 21; Length 840;
 Best Local Similarity 33.8%; Pred. No. 5,5e-107;
 Matches 291; Conservative 151; Mismatches 339; Indels 79; Gaps 18;

OY 12 WALLHPTGAPLCLUSQOL-----RMKGDIYVGLGLPFL-GEAEERAGURS 53

DB 4 WA-----AHLLSLQLYVCMAPSCQRTSESSPGRSLRDPDLFGLSLHSDC-----LQV 52
 OY 54 RTRDSSPVCTR---FSSNGLLALAMKMAVEINKNDLLPGLRIGYDLPTCSEPVAM 110
 DB 53 RHRPLVTSQDRPDSEFNGHYHFLQMRFTVEINNSGLLPNITLIGELAYDVCSES-ANV 111
 OY 111 KPSLMFLAKAGSRDIAACNCTQVQPRVLAIVIGPSSSELAWTGKFFSEFPLMPQVSGAS 170
 DB 112 YATLKVTLALQGRHTEIOKDLRNHSSKVAFIPGPDNTHAAYTAAALPDLPLVSYRAS 171
 OY 171 MELLSARETFPFPFRTVPSDRVQLTAAAEELLQEFGMNVAALSDDEXRGOLSTIFSALA 230
 DB 172 SVVLSAKRRKFPSEFLTFVSDRHQVEVMVQLQSFQFWMVVISLIGSGDYGQGLVQALBELA 231
 OY 231 AARGICIAHEGLVPLPRADSRIG--KVQDYLHOVNOSSVOVILLFASVHAHALEFNYSI 288
 DB 232 VPRGICVAFKDIVP-----SARVGDPRMQSMQHLAQARTVYVYFSNRHLARFFRSVY 287
 OY 289 SSRLSPKVWVASEAMLTSDVLMGLPQMAQMGTVLGFQRGAO---LHBPQ-YVKTHTAL 344
 DB 288 LNLTKGVWVASEDMALISTYITISVIGIGIVLGVAVQQRVGLKFEESYVAAYTA 347
 OY 345 AT---DPAFCALGEREGLDEDVVGRCPODCITLQNVS--AGLNHHQTFSVYAAYS 399
 DB 348 PSACEGSMCS-----TNQLCRCHFTTNNMPLTIGAFMSAAVRYVEAYYA 394
 OY 400 VAQALHNTLQCNASGCPADPKVQQLLENMNTLFHVGLPLRPDSSGNDMEYDLKLM 459
 DB 395 VAHGHQHLQIGTSEIC-SRGPVYPWQLIQLYKVFLLHENTVAVDDNGDTLGYDDIYAM 453
 OY 460 VMQGSVPLRHVGRNQS---LTERLKTIRNHTSDNOKPVSRCSRCQCEGQVRYKGPFS 516
 DB 454 DMNGPEMTFEITIGASLSPVHLIDINKTKIOMHGNNOVPASVCTDCLAGHRRVYVSGH 513
 OY 517 CCYDCVDEAGSYRONPDDIACTEGODEWSPERSTRCFRRRSRLANGEPAYVILL 576
 DB 514 COFECVPCFAGFLNMSLHLIQPCGTEWMAKRESTCFPRVEHLAMHEPISLVLIAN 573
 OY 577 SLALGLVLAALGLFVHNHDSPLVQASGCPPLACFGVLCGLVCLSTVLLPQGPSAPARCLAQ 636
 DB 574 TLLLLLVGTAGLFAMHHTPVRSAGRLCFMLGSLVPAVSCSPYSFGEPTAPACLR 633
 OY 637 QPISHLPTGCLSTFLQAAEFV---ESELPL---SMADRLSCLSGPMAMVLVLLAM 689
 DB 634 QPISLGFAPLPSCLTISFQVLITFKSTKVPPTFRYKMA-----QNHGAGLFVIYSS 686
 OY 690 LVEVALCTWYLVAFPPEVYVTDHMLPTBALVHCTRWSVSGLAHATNATLAFICLIGTF 749
 DB 687 TVHLILCTLWLVMTPRPTREYQRPHLVILECTEVNSVGFLLAFTNILLISTFVCSY 746
 OY 750 LVRSQPGCYNARAGITFMAVFTVWSFVPLNAVQVLLPRAYOMGALLCYGLIAAF 809
 DB 747 LKELPENYINAKCYTFELNLNFVSIWAFFTWASLYQSYLPAVNAVLAGLTLTLGSGFGSY 806
 OY 810 HLPRCYILMRQGLNTEPEF 829
 DB 807 FLPKCYVILCRPELNTEHF 826

RESULT 14
 ID AAY45021 standard; Protein: 840 AA.

AC AAY45021;
 DT 31-MAY-2000 (first entry)

DE Rat sensory transduction G-protein coupled receptor-B3.
 XX
 XX Rat; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
 KW sensory cell; taste receptor cell; screen; taste modulator;
 KW pharmaceutical; food industry; taste topographic map; tongue.

XX OS Rattus sp. Location/Qualifiers
 XX Key 1..580
 XX Domain /label= Extracellular_domain
 XX FT Domain
 XX WO200006592-A1.
 XX PD 10-FEB-2000.
 XX PF 27-JUL-1999; 99WO-US17099.
 XX PR 28-JUL-1998; 98US-0094465.
 XX PA (REGC) UNIV CALIFORNIA.
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Zuker CS, Adler JE, Lindemeier J, Ryba N, Hoon M;
 XX DR WPI; 2000-205451/18.
 XX DR N-PSDB; AA250743.
 XX PT New isolated sensory transduction G-protein coupled receptor, useful
 XX PT for developing products for use in studying and modulating the taste
 XX PT transduction pathway -
 XX PS Claim 22; Page 75; 83pp; English.
 XX CC The present sequence is a taste cell specific G-protein
 XX CC coupled receptor, GPCR-B3 which is involved in sensory transduction.
 XX CC This sequence was isolated from the 12AP rat circumvallate cDNA library.
 XX CC GPCR-B3 is specifically expressed in foliate and fungiform cells, with
 XX CC lower expression in circumvallate taste receptor cells of the tongue.
 XX CC The protein has an extracellular domain, seven transmembrane
 XX CC domains and an intracellular domain.
 XX CC The present sequence is used to screen compounds that modulate sensory
 XX CC signalling in taste cells, especially taste modulators useful in
 XX CC pharmaceutical and food industries to customise taste. The sequence
 XX CC can also be used as a probe for identifying taste cells and
 XX CC subsets of taste receptor cells such as foliate, fungiform and
 XX CC circumvallate. Such probes are also useful to generate taste
 XX CC topographic maps that elucidate the relationship between the taste
 XX CC cells of the tongue and sensory neurons leading to taste centres
 XX CC in the brain.
 XX CC
 XX SQ Sequence 840 AA;
 Query Match 26.7%; Score 1208.5; DB 21; Length 840;
 Best Local Similarity 33.8%; Pred. No. 6.9e-107;
 Matches 291; Conservative 151; Mismatches 339; Indels 79; Gaps 18;
 12 WALLHPTGAPLCSQOL-----RMKGDVVGGLPPL-GEAEAGIRG 53
 4 WA-----AHLISLQLVCMASFSCORTSSPEFSILPQDPLAGLSLHSDC---LQV 52
 54 RTRESSPVCTG---FSSNGILMALAMKAVEINNKSDLLPGLRIQYDLPTCSEPVVAM 110
 53 RHRPLVTSRCDPDSFNGHGLFQAMFTVEINNSALLPNTLIGELDYVCSG--ANY 111
 111 KPSLMFLAKGSRDIAYCINTQYQPRYLAVIGPHSSELMVTKFSSFLMPQVSGAS 170
 112 YATIRVLAOGPRHIELOKDLRNHSSKVAFIGPDNTHAVTFAALLGPFILMPLVSEAS 171
 171 MELLSARETFSPFRVPSRVOLTAALLOERGMVMVAALGSDDEYRGQGISIFALA 230
 172 SVVLSARKRPSFLRTPSDRHOVEVWQLQSFQWMSLIGSYGQGLGVQALIELA 231
 231 AARGCIHAGLVPLPRADSLRG--KVQDYLQOVNOSVQVVLVFAVNAHALFYNSI 288
 232 VPRGICVAFDIYVF-----SARVGDPRMQSMOHLAQARTTVVVFENRHLARFFESV 287
 289 SSRSPKVVWVASEAMLTSDLVNGLPGMAQMGVYIGLQGAQ---LHEPPO-YVKTHIAL 344

DB 288 LANLIGKVVWVASEDMALSTYITSTVGIQIGTVGVAVNOQROVGLKEFESYRAVATAA 347
 QY 345 AT---DPAFCSALGERDQGLEEDVVGORPCQDCITLQNS--AGLHHQFSSYAAVVS 399
 DB 348 PSACEGSMCS-----TNQCRCHFTTRNNPTLGAFSMAIRYEVAYTA 394
 QY 400 VAQALHNTLQCNAGCPADPVKPMOLLNNYNTLTFVYGLPLRFDSGGVMDYDKLV 459
 DB 395 VAHGLHQLIGTSEIC--SRGPVYPMQLQIQYKNFLHENTVAFDNGDITGLYDITAA 453
 QY 460 VWQGSVRLHDVGRFNS--LRTERLKIRNHSNDQKPVSRKQCEGVRRVYKGFHS 516
 DB 454 DWNPEWTFELIGSASLSPVHLIDINKTKIOWHGKNNQVPSVCTDIDLAGHRRVVGSH 513
 QY 517 CCYQVQCEAGSYRQNDLQACFCGODEMSPERSTCFRRRSFLAMGEPVALLLLL 576
 DB 514 CCEFCVPCENGFTLNMSELHICOPCGTEWAPKESTCFPTVTEFLAMHEPISLVLTAA 573
 QY 577 SLALGVLAALGLFVHHRDSPLVQASGGLACFGLVCLGLVCLSVLLPQGPSPARCLAQ 636
 DB 574 TLILLILVGTAGLFAWHFHTPVVRSAGRLCTLMGSLVAGSCSFYFPGEPVPAQLLR 633
 QY 637 QPLSHPLVTCGLSTFLQAAETV-----ESLPL---SMADRLSGCLRGPMALVLIAM 689
 DB 634 QPLFSLGFALFLSCLTIRSFOLVIFKFSKVPFRTWA-----QNHGAGLFIYSS 686
 QY 690 LVEVALCTWLVAFPREVVDHMLPTEALVHCTRSMVSGCLAHANNTALFICFCTF 749
 DB 687 TWHLLCTITWLVMTPTPTREYQRFPLVILECTEVASVGLFTNILLISTFVCSY 746
 QY 750 LVRSQPCYNRARGLAFMLAYFTWVSFVPLANVOVLRPAVOMKALLCYGLIAAF 809
 DB 747 LKRLPENYVNAKCVTSLINLEFWSIAFTMASIYGSVLPVAVNLAAGLTLISGFSGY 806
 QY 810 HLPKCYLLMPPGLNTPPEF 829
 DB 807 FLPKCYVILCRPELNTEHF 826
 RESULT 15
 AAY45026
 ID AAY45026 standard; Protein; 840 AA.
 AC AAY45026;
 XX
 DE 31-MAY-2000 (first entry)
 XX
 DE Rat sensory transduction G-protein coupled receptor-B3 variant #1.
 XX
 KW Rat; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
 KW sensory cell; taste receptor cell; screen; taste modulator;
 KW pharmaceutical; food industry; taste topographic map; tongue; variant.
 XX
 OS Rattus sp.
 XX Synthetic.
 XX Key Location/Qualifiers
 XX MISC-difference 33 /note="Wild type Ieu substituted by Ile"
 XX WO200006592-A1.
 XX PD 10-FEB-2000.
 XX PF 27-JUL-1999; 99WO-US17099.
 XX PR 28-JUL-1998; 98US-0094465.
 XX PA (REGC) UNIV CALIFORNIA.
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Zuker CS, Adler JE, Lindemeier J, Ryba N, Hoon M;


```

xx  WPI: 2000-205451/18.
DR
xx  New isolated sensory transduction G-protein coupled receptor, useful
PT  for developing products for use in studying and modulating the taste
PT  transduction pathway
PS  Disclosure: Page -, 83pp; English.
xx
CC  The present sequence is a polymorphic variant #1 of rat taste cell
CC  specific G-protein coupled receptor, GPCR-B3 which is involved in
CC  sensory transduction. GPCR-B3 is specifically expressed in foliate
CC  and fungiform cells, with lower expression in circumvallate taste
CC  receptor cells of the tongue. The protein has an intracellular domain,
CC  seven transmembrane domains and an intracellular domain.
CC  The GPCR-B3 sequence is used to screen compounds that modulate sensory
CC  signalling in taste cells, especially taste modulators useful in
CC  pharmaceutical and food industries to customise taste. The sequence
CC  can also be used as probe for identifying taste cells and
CC  subsets of taste receptor cells such as foliate, fungiform and
CC  circumvallate. Such probes are also useful to generate taste
CC  topographic maps that elucidate the relationship between the taste
CC  cells of the tongue and sensory neurons leading to taste centres
CC  in the brain.
CC  Note: The present sequence is not given in the specification but is
CC  derived from rat GPCR-B3 sequence shown in page 75 (AA45021).
xx
SQ  Sequence 840 AA:

Query Match      26.7%; Score 1208; DB 21; Length 840;
Best Local Similarity 33.9%; Pred. No. 7.7e-107;
Matches 291; Conservative 151; Mismatches 342; Indels 74; Gaps 19;

Oy  6 VLGSL--WAL-----LHPTGAPLCLSQGLRMKGDYVGLFPL-GEAEAGLRST 55
Db  9 LLSLQLYVCWASCSQRTSSPFGSIP-----GDFLAGLFLSHGDC---LQVRH 54
Oy  56 RPSFVCTR--FSSNGLLMALMAKAVEEINNKSDLLPGLRIAGYDLFDTCSEPVYAMKP 112
Db  55 RPLVTSQDRPDSFNGHGLFQAMRFTVEEINSSALLPNTILGLEYDVCS-ANYVA 113
Oy  113 SLMTAKAGSRDIAYCNYQYQRYLAVIGPHSELMANVTGKFSFPLMPOVSYGASME 172
Db  114 TLRLVALQGPRIEIQKLRHSSKVAFAIGPDNDHVTALLGPFLMPLVSYEASSV 173
Oy  173 LLSARETSPSEFRVPSDROYOLTAAEILOERGMWVAALSGDDEYRGQLSIFSAIAA 232
Db  174 VLSARKKPPSFLRTVPSDHOEVNVOILQSGFWWISLISYGDYGOALAEELAVP 233
Oy  233 RGICIAHGLVPLPRADSRGLG--KVQDVLHQVNOSSVQVVLFFASVNAHALFNYSSIS 290
Db  234 RGICVAFDIYVF---SARVGDPRMQSMQMLAQARTTVVVFVSNRHLARVFFRSVILA 289
Oy  291 RLSPKVWVASEAWLTSVLVNGLPQMAQMGTYIGFLQGAQ--LHEPQ-YKTHALAT 346
Db  290 NLTGVWVAASEMAISTYITSYTGIGITVGVAVOQROVPGLEKEFESYRAVTAAPS 349
Oy  347 ---DPAFCSALGERQGLEEDVVGORPCOCITLQNV--AGLNHQTFSYAAVYSVA 401
Db  350 ACPESSWCS-----TNOLCRCHTFTTRNMPTLGAFAFMSAAVRYEAIVAYA 396
Oy  402 QALANTLQCNASGCPADPVKFWOLLENNYNTLFVVGGLPLRFDSGQVNDMEYDLKLMW 461
Db  397 HGLHQLGCTSEIC-SRGVYPMQLQIYKVNFLHENTVAFDNGDGLGYDIIAMDW 455
Oy  462 QGSVRLHDVGRFNGS---LRTERLKIRNHTSDNOKPVSRGSCQEGOVRRVKGPHSCC 518
Db  456 NGPEMTFELIGSASLSPVHLIDINKTKIQHGKNNQVPVSVCTDCLAGHFRVVGSHHC 515
Oy  519 YCVVCEAGSYRQNDPDLACFCGDEWSPERSTCFRRSRFLAMGEPAVLLLLLLSL 578
Db  516 FECVCEAGSTFLNMSELHICOPCGTEWAPKESTYCFPTVEFLAMHEPISLVLAANTL 575

```

```

Oy  579 ALGLVLAAGLFVHHRDSPLVOASGGLACFGVCLGVCLSVLLFPGQSPARCLAOOP 638
Db  576 LLLILVGTAGLFAMHFHTPVVRSAGRLCFLMLGSLVAGSCSFYFFGGEPTVPACLLAQ 635
Oy  639 LSHPLTGCLSTFLQAAEIFV---ESELPL---SWADRUSGCLRGFWAMLVYLLAMV 691
Db  636 LFSIGFAIFLSCLTIRSFQLVLIIFKFSYKVPFRTMA-----QNHGAGLFIYSSSTV 688
Oy  692 EVALCTWLVAFPPRYVDMHMLPTREALVHCRTSRWSFGLAHATNATLAFCLFGLTFLV 751
Db  689 HLLICLTVMWTPTPTREYQREPHLVILECTEVNSVGFLAFTNILLISTFVCSYLG 748
Oy  752 RSQPCYNRAGLFEAMLAYFTTWVSFVPLLANOVVLRPAVOMGALLCYGLIAAPHL 811
Db  749 KELPENYNKACVPTSSLNFTVSWIAFTTMASTYQGSYLPAVNVLAGLITLSGFSGYFL 808
Oy  812 PRCYLIMROPGINTPEFF 829
Db  809 PKCYVILCRPELINTNEHF 826

```

Search completed: May 19, 2003, 09:50:04
Job time : 42.3065 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:47:40 ; Search time 20.1657 Seconds

(without alignments)
4061.679 Million cell updates/sec

Title: US-09-927-315-15

Perfect score: 4524

Sequence: 1 MLCFVVLGLSLMALHPGSGT.....GPGDAQGNQDNGTNGNGHE 852

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3229	71.4	858	2	JC7683
2	1068.5	23.6	1078	2	UC6715
3	1060	23.4	1085	2	SA40476
4	1057.5	23.4	1088	2	B56715
5	1054	23.3	1079	2	159362
6	670.5	14.8	872	2	JH0561
7	631.5	14.0	871	2	A46742
8	619.5	13.7	872	2	JH0563
9	612.5	13.5	1218	2	571376
10	604.5	13.4	915	2	A49874
11	597.5	13.2	879	2	JH0562
12	593.5	13.1	908	2	149142
13	592.5	13.1	879	2	JC7160
14	584	12.9	1171	2	A42916
15	580.5	12.8	1180	2	JC2132
16	580.5	12.8	1212	2	JC2131
17	574.5	12.7	1199	2	A41939
18	474	10.5	999	2	T27628
19	436.5	9.6	1267	2	T21340
20	323	7.1	551	2	T30806
21	240	5.3	962	2	D86186
22	238	5.3	976	2	T51137
23	237.5	5.2	925	2	T06128
24	236.5	5.2	941	2	T51135
25	221.5	4.9	953	2	E84732
26	220.5	4.9	1039	2	T45779
27	219	4.8	951	2	T51132
28	216.5	4.8	925	2	T51133
29	215.5	4.8	933	2	C96495

30	213	4.7	950	2	T51134
31	213	4.7	925	2	A84550
32	211	4.7	923	2	F84732
33	210	4.6	912	2	T51131
34	210	4.6	921	2	T51136
35	208.5	4.6	1099	2	T16283
36	205	4.5	960	2	JE0356
37	175	3.9	1103	2	JC5581
38	159.5	3.5	1108	2	T59385
39	151.5	3.3	1108	2	B55915
40	140	3.1	965	2	T51244
41	139.5	3.1	920	2	B84640
42	135.5	3.0	1323	2	T78557
43	133.5	3.0	1356	1	C45219
44	130.5	2.9	464	2	S75362
45	130	2.9	938	2	A46612

ALIGNMENTS

RESULT 1

JC7683

taste receptor T1R3 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001

C:Accession: JC7683

R:Kitagawa, M.; Kusakabe, Y.; Miura, H.; Ninomiya, Y.; Hino, A.

Biochem. Biophys. Res. Commun. 283, 236-242, 2001

A:Title: Molecular genetic identification of a candidate receptor gene for sweet taste

A:Reference number: JC7683; MUID:21222875; PMID:11322794

A:Contents: tongue

A:Accession: JC7683

A:Molecule type: mRNA

A:Residues: 1-858 <KIT>

A:Cross-References: DDBJ:AB049994

C:Comment: This protein, a seven-transmembrane receptor, serves as a mediator of the

A:Gene: t1r3

A:Map position: 4

C:Keywords: transmembrane protein

Query Match

Best Local Similarity 71.4%; Score 3229; DB 2; Length 858;

Matches 621; Conservative 67; Mismatches 147; Indels 22; Gaps 4;

QY	5	AVGLSLMALHPGTCAPICLSQQLRMKGDYVGLFPGAEBAEGLRSRTPSSPVCTR	64
DB	5	ATMGSLIAFLFELGMSASICLSQOFKAGDYLIGLFLGSTBEATLNORTQPNSTPCNR	64
QY	65	FSNGLTALAKMAVEEINNKSDLLPGRLGYDLFDCSEFVVMKPSLMLFARASRD	124
DB	65	FSPVLGFLMAKMAVEEINNGSALLPGRLGYDLFDTSEFVVMKPSLMLFARASRD	124
QY	125	IAAYCNITQYQPRVLAIVIGPHSSSELAIVGKFFSEFLMPQVSYGASMEILASRETPSFF	184
DB	125	IAAYCNITQYQPRVLAIVIGPHSSSELAIVGKFFSEFLMPQVSYGASMEILASRETPSFF	184
QY	185	RTVPSDRVOLTAABELLOEFGNMYAALGSDDEYRGQGISFSAALAAANGICIAHGLVP	244
DB	185	RTVPSDRVOLTAABELLOEFGNMYAALGSDDEYRGQGISFSAALAAANGICIAHGLVP	244
QY	245	LRADDSRLRGKYQVDLVHVOSSVOVVLFAVHAHALFNYSISRSRSPKWWVASEML	304
DB	245	QMDTSGQQLGKVLVDVLRQVNSKVQVYVFAARAVYSLEFSIHGGLSPKWWVASEML	304
QY	305	TSDLVWGLGMAQMGTVLGLTQGAQLHEFPQYVTKHLATADPAFCASALGEREGLED	364
DB	305	TSDLVWGLGMAQMGTVLGLTQGAQLHEFPQYVTKHLATADPAFCASALGEREGLED	364
QY	365	VVGQRCPCQDCITTLQNVASGL-----NHQTSVYAAYVSVAQALHNTLQCAAGCP	416
DB	364	VVGQRCPCQDCITTLQNVASGL-----NHQTSVYAAYVSVAQALHNTLQCAAGCP	416

Db 763 ITCHEGSLMALGFLIGTCLLAACFFFAFKSRKLPENFNEAKETIFSMLEFIWISPI 822
QY 780 PLLANQVVLRAVOMGALLCYLGILAFHLPKRCYLMKROGLNTP 827
Db 823 PAVASTYGFVSAVEVIALAASFGLLACIFENKYYIILFKPSRNTIE 870

RESULT 3
SA0476
Ca(2+) -sensing receptor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Nov-1999
C:Accession: SA0476
R:Brown, E.M.; Gamba, G.; Riccardi, D.; Lombardi, M.; Butters, R.; Kifer, O.; Sun, A.; H
Nature 366, 575-580, 1993
A:Title: Cloning and characterization of an extracellular Ca(2+)-sensing receptor from h
A:Reference number: SA0476; MUID:9407182; PMID:8255296
A:Accession: SA0476
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1085 <BR0>
A:Cross-references: GB:S67307; NID:9453108; PID:AA29171.1; PID:9453109

Query Match 23.4%; Score 1060; DB 2; Length 1085;
Best Local Similarity 30.0%; Pred. No. 6,7e-75;
Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

QY 6 VIGLSLMLHRTGAPLCLSQLRMKGDYVIGLFPPL-GEA-EEAGLRSTRPSSPYC 62
Db 10 LIAFTWCTSAVGP-----DQRAKKGDIILGGLFPHFGVAVKQDOLKS--RESVDC 61
QY 63 TPSSNGILMALAMKAVEEINNKSDLPGLRGLGDLFDTCSEPVYAMKPSMLFLA--KA 120
Db 62 IYNNRGRFRLQAMIFALIEINSSPALLPNNMLGRIPTCTMTVSALATLSFAVQNKI 121
QY 121 GSRDIAVANTYQYQYPRVLAVIGPHSSSELAMVTKFFSFFLMQVSYGASMEILLSARETF 180
Db 122 DSIINDEFNCSEHPISTIAVAVGATGSGISTAVANLLGLFIPOVYASASSRLSNKQF 181
QY 181 PSFRTVPDRYQVLTAAELLOEFGMMVVAALGSDDEXROGLSTFSALAAAGICIAHE 240
Db 182 KSFLLTIPNDEHQATAMADIIIEFRMNVGTIAADDOYRGIGKEFREAEEDIDIFS 241
QY 241 GLVPLPRADSRIGKVQDYLHQVNOSSVQVYVLLFASVHAHALFNYSISSRLSPKVVAS 300
Db 242 ELI-----SQYSDBEKIQVVEVIONSTAKVIYVSSGDLPLELKEIVARNITGRWLAS 297
QY 301 EWLITSDLVMLPGMAQMGTVIGFLQGAQLHEFPQYV-KTHIALATDPAFCSAL----- 354
Db 298 EAMASSSLIAMPDEYHVGVTIGFGLKAGQIPGFRFLOKVRPKSVHNGFAKEFEETE 357
QY 355 -----GEREGGL-----BEDVVGQRCPOCDCTTGN 380
Db 358 NCHLOEGAKGPLVDFTFLGHEEGGAKLSNSPTAFRPLCTGEBNISSTVETPYMDYTHLR- 416
QY 381 VSAGINHQTFSYVAAYVSAQALHNTLOC-----NASGCPADDPYKPMQLENNMYN 432
Db 417 -----ISYVAVLYASIAHALODIYTCIPGRGLFTNGS--CADIKKVEAMQVYLKHLH 467
QY 433 LFFHYG-GLPLRDSGANVMEYDLKL--VMQGSVPLHDVGRFN-GSLTERL----- 483
Db 468 LNFSTNMGOVTFDECGDLAGNYSTIIMHLSPEDSISV-FKEVGYVNAVYAKGERLFIN 526
QY 484 -KIRNHTSDNOKPVSRCRQCEGOVRR-VKGFHSCCYVDCVCEAGSYQNDDIACITFC 541
Db 527 EKILMSGREVPFNSCSDCLAGTRKGIITBEPCTCCFCVBCPODETSDETDASCDC 586
QY 542 GODEMSPERSTCFRRSRHFLAMGEPAVILLILLSLAGVLAALGLFVHHRDPSLYQA 601
Db 587 PDDFMSNHNHTSCIAKEIEFLSWTERFGIALTLFVAVLGIFLAFVIGVYIKRRNPIYVA 646
QY 602 SGGPLA---CFGLVCLGLVCL-SVLLFGQPSPARCLAQOPLSHLPLTGCLSTLEFLQAAE 657

Db 647 TNRELSTYLLFLSLC-----CFSSSLFEIGEPQDMTCRLRQPAFGISFVLCISCLIVKTN 702
QY 658 IFV--ESELPLS-----WADRISGCLRGWAMLVILLAMVEVALCTWYLVAPPEVYD 710
Db 703 VLLVEPAKIPTSRHRKMGOLNFO-----FLVFLCFCFMQVICALMLNPAFPSSYRN 754
QY 711 WHMLPTEAL-VHCRTSRWSFGLAHATNATNLATLFCIGTFVRSOPGCRNRAAGLEFAL 769
Db 755 -HELEDEIIFITCHEGSLMALGFLIGTCLLAACFFFAFKSRKLPENFNEAKETIFFSML 813
QY 770 AYTFTVSWEPPLANQVVLRAVOMGALLCYLGILAFHLPKRCYLMKROGLNTP 827
Db 814 IFEYIWSIPPAVASTYGFVSAVEVIALAASFGLLACIFENKYYIILFKPSRNTIE 871

RESULT 4
B56715
calcium receptor (clone pHpCar-5.2) - human
C:Species: Homo sapiens (man)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 21-Jul-2000
C:Accession: B56715
R:Garret, J.E.; Capuano, I.V.; Hammerland, L.G.; Hung, B.C.P.; Brown, E.M.; Hebert,
J. Biol. Chem. 270, 12919-12925, 1995
A:Title: Molecular cloning and functional expression of human parathyroid calcium rec
A:Reference number: A56715; MUID:95279439; PMID:7759551
A:Accession: B56715
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1088 <GAR>
A:Cross-references: GB:U20760; NID:9683746; PID:AAA6504.1; PID:9683747
C:Keywords: glycoprotein; receptor; transmembrane protein

Query Match 23.4%; Score 1057.5; DB 2; Length 1088;
Best Local Similarity 30.2%; Pred. No. 1.1e-74;
Matches 271; Conservative 155; Mismatches 365; Indels 107; Gaps 27;

QY 12 WALL-----HPGTGAPLCLSQLRMKGDYVIGLFPPLG---EEAEGLRSTRPSSPYC 64
Db 8 WVLLALTWHTSAVGP---DQRAKKGDIILGGLFPHFGVAVKQDOLKS--RESVDC 62
QY 65 FSSNGILMALAMKAVEEINNKSDLPGLRGLGDLFDTCSEPVYAMKPSMLFLA--KAGS 122
Db 63 YNFRGFRMLQAMIFALIEINSSPALLPNNMLGRIPTCTMTVSALATLSFAVQNKIDS 122
QY 123 RQIAVANTYQYQYPRVLAVIGPHSSSELAMVTKFFSFFLMQVSYGASMEILLSARETFS 182
Db 123 LNIDEFNCSEHPISTIAVAVGATGSGSTAVANLLGLFIPOVYASASSRLSNKQFNS 182
QY 183 FRTVPSDVPQVLTAAELLOEFGMMVVAALGSDDEXROGLSTFSALAAAGICIAHEGL 242
Db 183 FLRTIPNDEHQATAMADIIIEFRMNVGTIAADDOYRGIGKEFREAEEDIDIFSRL 242
QY 243 VPLPRADSRIGKVQDYLHQVNOSSVQVYVLLFASVHAHALFNYSISSRLSPKVVAS 302
Db 243 I-----SQYSDBEEDIHVVEVIONSTAKVIYVSSGDLPLELKEIVARNITGRWLASA 298
QY 303 WITSIDLVMGLPGMAQMGTVIGFLQGAQLHEFPQYV-KTHIALATDPACSAALGEE--- 358
Db 299 WASSSLIAMPQYFHVVGVTIGFGLKAGQIPGFRFLOKVRPKSVHNGFAKEFEWETFC 358
QY 359 -----QGLEEDVVGQRCPOCD-----CITLQVNSA-----GLNHHO- 389
Db 359 HLOEGAKGPLVDFTFLRGHEES--GDRFNSSTAFRPLCTGEBNISSTVETPYIDYTHLR 416
QY 390 TFSYVAAYVSAQALHNTLOC-----NASGCPADDPYKPMQLENNMYNLTFF-HVGL 440
Db 417 SYNVAVLYASIAHALODIYTCIPGRGLFTNGS--CADIKKVEAMQVYLKHLHNFNTNMGE 475
QY 441 PLRFDSSGANVMEYDLKL--VMQGSVPLHDVGRFN-GSLTERL-----KIRNH--- 488
Db 476 QVTFDECGDLGVNYSSTIIMHLSPEDSISV-FKEVGYVNAVYAKGERLFIN 526
QY 489 -----TSDNOKPVSRCRQCEGOVRR-VKGFHSCCYVDCVCEAGSYQNDDIACITFC 540

Db 535 REPTLVSVLQVPPSSNCRCLAGTRKGIIEGPTCCFEEVCECDGSEYSDTASAKNK 594
 QY 541 CGODEMSERSTRCFRARRSLFAMCEPAVLLILLSTALGLVLAALFLPHHRDSPVQ 600
 Db 595 CPDDFMSMENTSCIAKEIFELSTWEPFGIALTLFVAVGILFVAVLVEFKFRPTLYK 654
 QY 601 ASGGPLA---CGFLVGLVCL-VYLPPGQSPARCLAOPLSHPLTGLSTLFLQAA 656
 Db 655 ATNRELSTLLFLSILC-----CFSSSLFFIGEPQDWTCRLRPAGFISVLCILVATN 710
 QY 657 EIFV--ESELPLS-----WADRLSGCLRGPMANLVLLAMVEVALCTWYLAPEPVY 709
 Db 711 RVLVFEAKIPTSFHRKMWGLNQ-----FLVFLCTEMQIYICVIMLYTAPSSYR 762
 QY 710 DMHMLPTALVHCRTSRVSWFGLAHATNATLFLCFLGTFVRSOPGCYNRARGTLFAML 769
 Db 763 NOLEDELIFITTCHEGSLMALGFLIGYTCLLAALCFEFAFNRKLPENFENAKFTTFSML 822
 QY 770 AFTTWSFVPLANVOVYLRPAVOMGALLLCVIGILAFAHLPRCYLLMRPGILNTPPE 827
 Db 823 IFFIYWFISFIPAYASTYKGFVSAVEVIALASFGLLACIFENKLYIILFRPSRTIE 880

RESULT 5

159362
 calcium/polyvalent cation-sensing receptor precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 02-Jul-1996 #sequence-revision 02-Jul-1996 #text-change 05-Nov-1999
 C:Accession: 159362; A55594
 R:Ruat, M.; Molliver, M.E.; Snowman, A.M.; Snyder, S.H.
 Proc. Natl. Acad. Sci. U.S.A. 92, 3161-3165, 1995
 A:Title: Calcium sensing receptor: molecular cloning in rat and localization to nerve terminal
 A:Reference number: 159362; MUID:95241465; PMID:7724534
 A:Accession: 159362
 A:Status: translated from GB/EMBL/DBDf
 A:Molecule type: mRNA
 A:Residues: 1-1079 <RES>
 A:Cross-references: EMBL:U02089; NID:9790578; PIDN:AAC52195.1; PID:9790579
 A:Experimental source: striatal
 R:Riccardi, D.; Park, J.; Lee, W.S.; Gamba, G.; Brown, E.M.; Hebert, S.C.
 Proc. Natl. Acad. Sci. U.S.A. 92, 131-135, 1995
 A:Title: Cloning and functional expression of a rat kidney extracellular calcium/polyvalent
 A:Reference number: A55594; MUID:95110508; PMID:7816802
 A:Accession: A55594
 A:Molecule type: mRNA
 A:Residues: 1-133, 'X', 135-1079 <RIC>
 A:Cross-references: GB:U10354
 A:Experimental source: kidney
 C:Keywords: calcium; glycoprotein; phosphoprotein; transmembrane protein
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:187-212/Region: hydrophobic
 F:613-635/Domain: transmembrane #status predicted <TM1>
 F:633-670/Domain: transmembrane #status predicted <TM2>
 F:683-700/Domain: transmembrane #status predicted <TM3>
 F:725-744/Domain: transmembrane #status predicted <TM4>
 F:770-790/Domain: transmembrane #status predicted <TM5>
 F:806-828/Domain: transmembrane #status predicted <TM6>
 F:841-860/Domain: transmembrane #status predicted <TM7>
 F:901-261,267,386,468,488,594,893,1005/Binding site: carboxylate (Asn) (covalent) #status
 F:794/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
 F:899,901/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted

Query Match 23.3%; Score 1054; DB 2; Length 1079;

Best Local Similarity 30.3%; Pred.No. 2e-74;

Matches 269; Conservative 163; Mismatches 363; Indels 92; Gaps 28;

QY 9 LSWALL-HPTGAPLCSQRLRMKGIVYLGCLPLG---BAERAGLSRTRPPSPVCTR 64
 Db 8 LALLALHSSAYGP---DQRAOKKGDILIGLPEPIHGVAKADQDLKS--RPESVEICR 62
 QY 65 FSSNGLLMALMKMAVEFINNKSDLLPGLRGYDLFDGCSPPVAMKRSMLFLA--KAGS 122

Db 63 YNFGFRLQAMIFAIEINSSPLLPMNTLGYRIPTDCTNTVSTALFATLSFAONKIDS 122
 QY 123 RDIAYCNYTOYQPRVLAVIGPHSSSLAMVYGFKEFSFELPQVSYGASMLLSARETPS 182
 Db 123 LINDFCSCSHITSTIAVATGSGYSTAVANLLGLFIYIQVYASSRLSNKNOKS 182
 QY 183 FFRVPSDRVOLTAABELLQFGNMVNAALGSDDEYGRGUSLIFSALAANGICIAHEGL 242
 Db 183 FLRTIPNDEHOTAMADIIIEYFRNMVGTIAADDYGRGIEKFRFEAEEDICIDISEL 242
 QY 243 VPLRADSRIGKYQVDVHOVNOSSQVYVLLFASVHAHALFENSSIRLSPKYVWSEA 302
 Db 243 I---SQYDEEELQVVEVIONSTAKYIVFSSGPDLEPIKEIVRNITIGRIWLSSEA 298
 QY 303 WLTSDLVMLGPMAGMGTVLGFLQGAOLHFPQYV-KTHALATDPAFCALGER---- 357
 Db 299 WASSSLIMPEYFHVYGGTIGFGLKAGIIPFREFLQVHNRKRYVHNFARFEEPEFNC 358
 QY 358 --EGLF-----EDVVGRCPCQD-----CITLQNVSA-----GLNHQ--TF 391
 Db 359 HLOGAKGPLVDITFVRSHEGGRNLLNSYAFRPLCTGDNINSVETPYMDYELRISY 418
 QY 392 SYTAAYSVAGALNLTLOC-----NAGCPRADDPYKPKQLLENMNLTF-HVGLPL 442
 Db 419 NVYLAIVSIALADIDYTCPLGRGLFTNGS--CADIKVEAMQVAKHLHLFTNNKGOV 477
 QY 443 RFDSSGNVMEYDKLW---VMQGSVPRLDVGRFN-GSLTERL-----KIRWHTSDNQ 493
 Db 478 TFDECGDLVGVYSIINMHLSPEDSIV-FKEVGYNYAKGERLFINEXELIWSGFSRE 536
 QY 494 KPVSRCSRCQCEGOVRR-VKGFHSCYDCVCEAGSTRQNDPIACTFCGODEMSPERST 552
 Db 537 VPFNSCSHDCAGYRKGITIEBEPCEFCVCCPGEYSGEFLDASACDKCPCPDFMSNHT 596
 QY 553 RCFRRSRFLAMGEPVAVLLILLSTALGLVLAALGLVHHRDSPVQASGPILA---CF 609
 Db 597 SCIAKEIFLWTEPFGIALTLFVAVLGLTFVAVGILFVAVLVEFKFRPTLYK 656
 QY 610 GLVCLGLVCL-VYLPPGQSPARCLAOPLSHPLTGLSTLFLQAAEIFV--ESELPL 666
 Db 657 SLIC-----CFSSSLFFIGEPQDWTCRLRPAGFISVLCILVAVLVEFKFRPT 712
 QY 667 S-----WADRLSGCLRGPMANLVLLAMVEVALCTWYLAPEPVYTDHMLPTAL-V 720
 Db 713 SFHRKMWGLNQ-----FLVFLCTEMQIICIIYLYTAPSSYRN-HELDEIIFI 763
 QY 721 HCRTRSWVSGFLAHATNATLFLCFLGTFVRSOPGCYNRARGTLFAMLAVITWSFVP 780
 Db 764 TCHGSLMALGSLIGYTCLLAALCFEFAFNRKLPENFENAKFTTFSMLIFFIYWFIS 823
 QY 781 LLANVOVYLRPAVOMGALLLCVIGILAFAHLPRCYLLMRPGILNTPPE 827
 Db 824 AYASVYKGFVSAVEVIALASFGLLACIFENKLYIILFRPSRTIE 870

RESULT 6

JH0561
 metabotropic glutamate receptor 2 precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Jun-1992 #sequence-revision 30-Jun-1992 #text-change 13-Sep-1998
 C:Accession: JH0561
 R:Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
 Neuron 8, 169-179, 1992
 A:Title: A family of metabotropic glutamate receptors
 A:Reference number: JH0561; MUID:92110002; PMID:1309649
 A:Accession: JH0561
 A:Molecule type: mRNA
 A:Residues: 1-872 <TMAN>
 A:Experimental source: brain
 C:Comment: This protein is coupled to a G protein and evokes a variety of functions b
 C:Superfamily: metabotropic glutamate receptor 4
 C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane p
 F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-872/Product: metabotropic glutamate receptor 2 #status predicted <MET>
F:568-550/Domain: transmembrane #status predicted <TM>
F:605-625/Domain: transmembrane #status predicted <TM>
F:637-655/Domain: transmembrane #status predicted <TM>
F:680-700/Domain: transmembrane #status predicted <TM>
F:726-747/Domain: transmembrane #status predicted <TM>
F:761-782/Domain: transmembrane #status predicted <TM>
F:795-819/Domain: transmembrane #status predicted <TM>
F:203,286,338,402,547/Binding site: carboxylate (Asn) (covalent) #status predicted
F:601,675,827,837,84/Binding site: phosphate (Ser) (covalent) #status predicted
F:832/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match	14.8%;	Score 670.5;	DB 2;	Length 872;
Best Local Similarity	27.8%;	Pred. No. 2.1e-44;		
Matches 256;	Conservative 133;	Mismatches 385;	Indels 147;	Gaps 37;

QY	1	MUGPVRVGLSLMALLHPTGAPICLSOOLRNKRGVYVIGGFPL-----GEAEAGLSRTR	56
Db	4	LLGLFAL--LLLMGVAABEPA-----KKVLTLEGLVJGFEVPHQKGGPABEGPVNEHR	57
QY	57	PSPBCTRFSSNGGLMAMKMAVEEINNNSDLLPGLRGLGDTEDTCSEPVAMKPSLME	116
Db	58	-----GIQRLMEAMFALDRINRPHLLPGVRLGAILDSCSKTHALEQALDF	105
QY	117	LAKA-----GSRDI---AAVCNTQYQPRVLAIVGHSSELMATVTKFFSFLMPOVST	167
Db	106	VRAISLSRGADSSRIICPDGSAVATSDAPRTAVTGYIGSISDVSIOVANLRLRPOIQISY	165
QY	168	GASMEILLSARETSPFFRFTVDSDEVOVTLAAELLQEEGMMVMVALGSDDEGROGLSIFS	227
Db	166	ASTAKSLDSKRDYVFPARTVPPDFQAKAAEILRFNTWYVTSVYABEGYGEIGIAFE	225
QY	228	ALAAARGICIAHEGLVPLPRADSRILGKVQDVLHQVNOSSVOVLLFASVAAHALENYS	287
Db	226	LEARRNRCVATSERKVGAMRAAFEGVYRALT---OKPSARVAVLTFRSEDARELL--A	280
QY	288	ISSRLSPK--VWVASEPMTISDLVNGLDGMQMGVULFQLRGOLHFFP--QYKTHLAL	344
Db	281	ATORLNASFTVASDGMGALDSVYAGSERRAEGAIT-----IELASTPLSDSFSTQSL	344
QY	345	-----ATDPAFCALSGEREGOLEDDVVGORPODCITLONVSAGLNHHOTFSYAAVYS	359
Db	335	DPWNNSRNPWF-----REFEERHFHCSFPRQDCAHSLRAVPREQESKIMF--VYNAVYA	387
QY	400	VQAOLHN---TLQCNASG--CPAODPVKFWOLLEN-KTNLTFHNGDL-----LREDSG	448
Db	388	MAHALHNNHRAICENTHTLDCAMPKPVNGRRLYKXQFVLNVKFDAPFRADTDDEVERDFRG	447
QY	449	NWDEYDLKTLVWOGSVP--RLHDVGRFNGSLRTEBLKIRHMT--SDNCKPVSRCSROOEG	506
Db	448	DGIRYINFTYLRRGSGRIYKQKGYABEGILTDTSTIPNAPSAGCLPASKRSEPLON	507
QY	507	QVRRYKGHSCCYDCVDCGASGRQNPNDIACFPQGDDEMSPEKSTRCFRRSRFTLWAG	566
Db	508	EKVSQPEVCEWCMICIPQPEYERL--DEFQADCGGLGYMPNMSLTCFELPQEYIMGD	565
QY	567	PAVLLLLLLSLAGVLAALGLEVHHRDSPLVOASGCPACBGLVCLGLVCL-----SVLL	623
Db	566	AMAYGPVITACLGALATLEVIGVFRNAPPVYKASRELL-CY--ILLGFGFLCYCKTFV	622
QY	624	FPQGPSPARCIAOQPLSHLPTGLSTPLFOA---AEIF-----VESELPLSWADLSG	674
Db	623	FIAPRSTAVCTLRLRLGIGTAFSVCYLSALLTKTNIAIIFGARGRGARQRPFIISASVAL	682
QY	675	CLRGPMANLVVLLAMEVALCTWYLVAFP-----PEVTTDMMHLPTEALVHCRT	724
Db	683	C-----LALISQOLLIVA--AMLVYEAQGTKETAPRERREYVT-----LRCHN	723
QY	725	RSWVSFGIAHNTNTIAFLCIGELFELVRSQPGCVNRRAGLTFMNLAFITWVSVPPLAN	784
Db	724	RDASMLG--SLAYNVLLALCTLYIAFKTRCKEENNENKAFIGETWYTTCTIIMLAFLE--	780
QY	785	VQVYVLRPVQMGALLC-----VLGIIAFAHRLRCVLYLMKROPGLN-----TPEF	828

Dd			:	:	:					:	:	:	:						
	791	--YYHSSDYRQVTTTMCYSVLSGSYVLGLFA--																	835
Oy	829	FLGGSPGDAGQGNDGNTGNQG																	849
			:	:	:					:	:	:	:						
Dd	836	-----GSAAPRASNLG-QG																	849
			:	:	:					:	:	:	:						

RESULT 7
A46742
metabotropic glutamate receptor, mGluR6 - rat

C:Species: *Rattus norvegicus* (Norway rat)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000
C:Accession: A46742
R:Nakajima, Y.; Iwakabe, H.; Akazawa, C.; Nawa, H.; Shigemoto, R.; Mizuno, N.; Nakani
J. Biol. Chem. 268, 11868-11873, 1993
A:Title: Molecular characterization of a novel retinal metabotropic glutamate recepto
A:Reference number: A46742; MUID:93280152; PMID:8389366
A:Accession: A46742
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-871 <NAK>
A:Cross-references: GB:DI3963; NID:g391856; PID:BA03066.1; PID:g391857
A:Experimental source: retina
A:Note: sequence extracted from NCBI backbone (NCBIN:133246, NCBIPI:133250)
C:Superfamily: metabotropic glutamate receptor 4
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match	14.08;	Score 631.5;	DB 2;	Length 871;
Best Local Similarity	25.68;	Pred. No. 2.5e-41;		
Matches 225; Conservative	146;	Mismatches 410;	Indels 99;	Gaps 31;

```

0Y 4 PAVGLSLMALLHPETGAPLCLSOOLRNKEDYVLGGLFPLGAEAEAGLRSTRFPSSPVCT 63
Db 5 PVLTLMLTAMWMLSOAGIA---CGAGSVYRLAGGLTLGGGLFPVHARGAG---RACG 52
0Y 64 RF-SSENGLLAMKMAVEEINNNSDLLPGRLGYDLPDPCSEPVYMAKRSIAME-----LA 118
Db 53 ALKKQGGHRLIEAMLYALDRYNADPELLPGRLARLIDLTCSRTYVLEOGLSVQALIR 112
0Y 119 KAGSHDIAYCNYTOYD-----RLVAVYGHSSSEIAYTGKFFSEFLIMPOVSYGASM 17U
Db 113 GRGDDEASVPCGCVPLRLSAPPERVAVAYGASASYSIMVAVNLYRLAIPOLSTASTA 17Z
0Y 172 ELSARETFPFSEFRVPSDRQLTAAAEELLDFEFGNNVYALGSDDEFGRGLSTFSALA- 230
Db 173 PELDSTRYDEFSRVYPPDSYQAOAMDYVRLAMVNYSTLASGNGEGSEVGEAFVQISR 232
0Y 231 AARGCIIAHEGIVLPYPRADDSRLKQVODYLVQVNOVS-SVOYVLLFASVHAHALFVNSIS 289
Db 233 EAGVCVIAOS--IKIPR--EKPGEFHKVIIRLMEIHPAKRITITFAEDDIRYLEATRO 288
0Y 290 SRLSPK-VWVASSEAMLT5-DLVMLPGMAQMGTVLGLFQRAQOLHEEPYUKTH----- 341
Db 289 ANLUGHFLWAGSDSMGSKISPLMLNEEA-VGALT-ILPKASIDGRDQFEMTSLENNRR 346
0Y 342 ----LALATDAF-----CSALGEBEGGLE-----EDVYGQRCPQDCITLTONV5AGLNNHQ 388
Db 347 RNIMFAEFMEENFNCKLITSSGGGDDSTRKCTGEBERIGD-----SAYEDEGK 394
0Y 390 TFSVYAAYVSYAQAOLHNTLQCNASG---CPAOPVYPMQOLEMYNLTFFH-VQGLPLRF 444
Db 395 VQVFDVAYVATLALHLSHQALCSGHTGLCSRAMERTQGRITLHITRAYVRFGSAGTYMF 454
0Y 445 DSSGNVMEYDL-KLWVQGSVPR-LHDVGRFGNSLTERELKIRMTSDNOKPVSCSR 501
Db 445 NENGDAPEGRYDIFQVATNGSAGSAGGYOAVQOMAEALRLDMEVLRWMSGDHEHVPSSCSL 514
0Y 502 QOGBQYVNR-VKGFHSCCYDCVDDEAGSYKRNPNPDIACTCGGDDEWSPPESTRKFRRSR 560
Db 515 PCGPERKMKVKG-VPCWHEACD--GYRPQVDEFTCEACPGIMRPTNHTGCRPTPV 571
561 FLAMEBPAVLLLLLSTALGVLAALGLFHNHNDSPVQVAGSGRLACFGLVGLGVLCS 620

```

Db 572 RLWSSPAAAPLLAVLAVGIMATTTTMAFEMRHNDPTIVRASGRELSTVLLTGIFLIYAI 631
 QY 621 VLLPGOPSPARCLAOCPPLPTGLCTSTFLQA--AEFVSESLPLSWADRLSGCLR 677
 Db 632 TELVAEBCAICARRLGLGTLTSLTSTLTNRIRYRFEQGRKRVTPPTLS---- 687
 QY 678 GPMAVLVYL--LAMLVEVALCTWYLVAFPPPEVTDHMLPT-----EALVHCRTSRWS 729
 Db 688 -PTSQVTFEGLTSLQVGVYIAM--LGAOPPHSVIDYEQRPVDEQARGVLKCDMSDL 745
 QY 730 FGLAHATNATLAFICFLGTFVLRSPGQCYNRRARGITFEMALYFTTWVSFVLLANV---- 785
 Db 746 IGCL-GTSLMLMTCTYTAIRARGVPEFENAKRPGTFMTTCITLWLFVIFFGTAQSA 804
 QY 786 -QVVLPRAVOMGALLCYLGIILAFHLPFCYLIMKROGLN 824
 Db 805 EKIIQTTLTVSLTSLASVSLGMLVYKTVILFHPEDN 844

RESULT 8

JH0563
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
 C:Accession: JH0563; I58149
 R:Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
 Neuron 8, 169-179, 1992
 A:Title: A family of metabotropic glutamate receptors.
 A:Reference number: JH0561; M01D:92110002; PMID:1309649
 A:Accession: JH0563
 A:Molecule type: mRNA
 A:Residues: 1-912 <TAN>
 A:Experimental source: brain
 R:O'Hara, P.J.; Sheppard, P.O.; Thøgersen, H.; Venezia, D.; Haldeman, B.A.; McGrane, V.;
 Neuron 11, 41-52, 1993
 A:Title: The ligand-binding domain in metabotropic glutamate receptors is related to bac
 A:Reference number: I58149; M01D:9333699; PMID:8338667
 A:Accession: I58149
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-123; 'R', 125-912 <RES>
 A:Cross-references: GB:M90518; NID:g205400; PIDN:AA93190.1; PID:g205401
 C:Comment: This protein is coupled to a G protein and evokes a variety of functions by m
 C:Genetics:
 A:Gene: GUR4
 C:Superfamily: metabotropic glutamate receptor 4
 C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; phospho
 F:1-32/Domain: signal sequence #status predicted <SIG>
 F:33-912/Product: metabotropic glutamate receptor 4 #status predicted <MET>
 F:588-610/Domain: transmembrane #status predicted <TRI>
 F:625-645/Domain: transmembrane #status predicted <TII>
 F:657-675/Domain: transmembrane #status predicted <III>
 F:700-720/Domain: transmembrane #status predicted <TIV>
 F:751-772/Domain: transmembrane #status predicted <TRV>
 F:786-807/Domain: transmembrane #status predicted <TVI>
 F:822-847/Domain: transmembrane #status predicted <VII>
 F:98,301,454,484,569/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:621,689,695,859,870/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 13.7%; Score 619.5; DB 2; Length 912;
 Best Local Similarity 24.6%; Pred. No. 2.3e-40;
 Matches 222; Conservative 145; Mismatches 397; Indels 137; Gaps 31;

QY 4 PAVGLSLMALHPGT-----GAPCLISQOLRMKGDVYLGPFL-GEAEFAGLRSTRP 57
 Db 15 PLCLLSLIXAWVSSLSCKPKGPHMS--IRIGDITLGLFVYHKGSE----- 63
 QY 58 SSPVCTRF-SSNGILMALMKMAVEEINKSDLPGLRLGYDLPTDCEPVPAMKPSLME 116
 Db 64 -GKACGELKKEKGIHRLKLEAMFLDRINNDPLPNTITLGRILIDTCSRDTHALEOSTLF 122
 QY 117 LAKAGSD-IAAYCN-----TYQPRVLAIVIGHSSELAMVTKGFFSFLMPQVSYGAS 170

Db 123 VQALIEKDGIEVRGSGGPPITRKPERVGVIGASGSVSIMVANIIRLFETIPQISYST 182
 QY 171 MELLSAREFPSPFRTPYSDRVLTAALAELOEGRMMVVALGSDDEYGRGLIFSA 230
 Db 183 APDLSNDRYDFEFRVYPSDTYQANAVDIYALKMNTVYSLASEGSGEYNAFQKS 242
 QY 231 AAR-GICIAHEGLVLPRAADSRIGKYQDVLAQ--VNQSSQVVLLEFAS-----VHAH 281
 Db 243 RENGVACIAQS--VKIRP--EPKTEGDKIKRILETSMNAGIITIFANEDDIRVLEAR 298
 QY 282 ALFKYSSISRLSPKVPVASEAM--LTSDLVYGLPMAQMGVYLGLOGAOLHEPQVYKT 340
 Db 299 R-----ANQGFHEFMWGSWSGSKSAPVLEEVAAE-GAAT-ILPRMSTRGDRYESS 350
 QY 341 H-----LALADPAFCSALGEREGLDEEDVVGORPCQDCTILQNVASGLNHHQ 389
 Db 351 RTLDNNRNRTMFAFEWEDNFKCL---SRNALKKGSHKKCTNERNIGQD--SAVEDEK 405
 QY 390 TFSYAAVYSVAQALHNTLQCNASG---CPADPVKPMOLLENMYLTFH-VGGLPLRF 444
 Db 406 VQFVIDAVYAMGHALHMHNDLCPRGVGLCPRMDPVDGTOLKXIRRVNFGSIAGNPFVTF 465
 QY 445 DSGNVDMEXYDLKLMVWQGSVPRLDVGRNGSLRTERLKRTRHTSNOKRVSRCRQC 504
 Db 466 NENGDAQRDYDIYQQLNNGSAEYKTVIGSWTDHLRLIERMOWGSGQQLRSTICSLPCQ 525
 QY 505 EGQVRR-VKGFHSCCYDCVDEAGSYRONPDDIACFCGODEMSPERSTRFRSRFEL 563
 Db 526 PGEKKVYKGM-ACWCNCEPRT--GYQYQYDRYCKKCPIDMRTEKNTSQPIPIYKLE 582
 QY 564 WGEPAVLLLLLSLALGLVLAALGLFVHNRDPLVQASGGPLACFGLVCLGLSVLL 623
 Db 583 WDSWAVVLPFLAVVAGIATLFFVVTFVRYNDPTIVRASGRELSTVLLAGIFLCYATFTFL 642
 QY 624 FPGQSPARCLAOCPPLHPLTGLSLTFLQAALF-----VSESLPLSWADRLSG 674
 Db 643 MIAPBDLQTCRLRIEYGLGMSISYALLTNTNIRYRFEQGRKRSVSAPEFISASOLA- 701
 QY 675 CLRGPMALVVLMLVEVALCTWYLVAFPPPEVTDHMLPTLALVHCRTSRWSFGLAH 734
 Db 702 -----IFILISLQLLGLGVFVVD-PESHVVDQDQDTLDRPAR-----GVLK 745
 QY 735 ATNATLAFCLGTFVY-----NSQPCYNARAGLTAMALAFYTWVSFVPL 782
 Db 746 CDISDLICLLGLYSMLMTCTYTAIRKRGVPEFENAKRPGTFMTTCITWLFATFIF 805
 QY 783 -----ANVOVVLPRAVOMGALLCYLGIILAFHLPFCYLIMKROGLNTP 826
 Db 806 FGTSQADKLYIQTTTLTVSVLSASVSLGML-----YMPKYITILFHPEDNP 854
 QY 827 E 827
 Db 855 K 855

RESULT 9

S71376
 C:Species: Oncorhynchus masou (cherry salmon)
 C>Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 01-May-1998
 C:Accession: S71376
 R:Kubokawa, K.; Miyashita, T.; Nagasawa, H.; Kubo, Y.
 FEBS Lett. 392, 71-76, 1996
 A:Title: Cloning and characterization of a bifunctional metabotropic receptor activat
 A:Reference number: S71376; M01D:96354880; PMID:8769318
 A:Accession: S71376
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-1218 <KUB>
 C:Keywords: glycoprotein; phosphoprotein
 F:603-625/Domain: transmembrane #status predicted <TM1>
 F:640-660/Domain: transmembrane #status predicted <TM2>

Query Match 13.7%; Score 619.5; DB 2; Length 912;
 Best Local Similarity 24.6%; Pred. No. 2.3e-40;
 Matches 222; Conservative 145; Mismatches 397; Indels 137; Gaps 31;

QY 4 PAVGLSLMALHPGT-----GAPCLISQOLRMKGDVYLGPFL-GEAEFAGLRSTRP 57
 Db 15 PLCLLSLIXAWVSSLSCKPKGPHMS--IRIGDITLGLFVYHKGSE----- 63
 QY 58 SSPVCTRF-SSNGILMALMKMAVEEINKSDLPGLRLGYDLPTDCEPVPAMKPSLME 116
 Db 64 -GKACGELKKEKGIHRLKLEAMFLDRINNDPLPNTITLGRILIDTCSRDTHALEOSTLF 122
 QY 117 LAKAGSD-IAAYCN-----TYQPRVLAIVIGHSSELAMVTKGFFSFLMPQVSYGAS 170

F:672-690/Domain: transmembrane #status predicted <TM3>
 F:717-737/Domain: transmembrane #status predicted <TM4>
 F:761-782/Domain: transmembrane #status predicted <TM5>
 F:786-817/Domain: transmembrane #status predicted <TM6>
 F:826-850/Domain: transmembrane #status predicted <TM7>
 F:104,233,403,525,757/Binding site: carbohydrate (asn) (covalent) #status predicted
 F:656,659,961/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status pre
 F:705/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:892/Binding site: phosphate (Ser) (covalent) (by CAM-dependent kinase) #status predic

Query Match 13.5%; Score 612.5; DB 2; Length 1218;

Best Local Similarity 25.7%; Pred. No. 1.2e-39;
 Matches 226; Conservative 140; Mismatches 370; Indels 143; Gaps 34;

QY 30 RMKGIVYVGGLEPL-----GEAEAGLRSTRSPSSVCTRSSNGILMLA 75
 DB 46 RMDGIIIGALFVHHQDSTQVARKCKGEVRE-----OYGIQRYEA 87
 QY 76 MKMAVEEINKSDLLPGRLGYDLFDTCSEPVAMKPSLMFLA-----KAGSRD 124
 DB 88 MFHILDRINSNPILLPTTLCCEIRDSQWSSVALLQSEIRDSLSIRDDDKDGTSR 147
 QY 125 IAAVCNTQYQYR-----RYLAVIGPSSSELMATGKFFSFLMPQVSYGASMEILLARET 179
 DB 148 QMCIEGMPSPQPAKTRPIAGVIGPSSSVAIQVONLQLFNIPQIAYSATSIDSLDKTL 207
 QY 180 FPFSEFRVPSRVQLTAAELLQEGWVVAALGSDDEYQOGSLISFALAARQICIAH 239
 DB 208 FKFLRVVPSDTLQARLALDVKRNMVTVSAVHTGEGSGMAFELASQEBELCIAH 267
 QY 240 EGIPLVPRAD--DSRLKQVDVLHQVQSSVQVLLFASVHAALNFYSISRLSPKV 296
 DB 268 SDKIYSNNGEKHFDBLLKRLERL-----PKARVVVCECEGTIVAGLL--MAMR---RL 316
 QY 297 WVASEAWLTSLDWGLPFGMAQGYVL-GFLQRG-----AQDH---EFPQY-VKTHIAL 344
 DB 317 GVAGEF-----LLIGSDGMADRDEVEEGYEQEAVGIVTKLHSEVTSFDFYFLKRLNLT 371
 QY 345 AT-DPAFCASALGERO-----GLEEDVVGQRCPQDCIT---LQNVASGLNHHTFESVY 394
 DB 372 NTRNPFEPFQOHROCIPIGHPLENNMYKRNKSGYESLEDNYVDOSKMGF-----VI 424
 QY 395 AAVYSVAQALHNTLQCNASG---CPAADPYKPMQLLENMTNLTP-HVGGPLRLPDESSGN 449
 DB 425 NAITAMAGLDMHSLCPGHVGLCKAMPIDGSQLLEFLMRTSTGYSGDEWWDENGD 484
 QY 450 VDMETDLK-LMWOGSVPLRHLDVGRFNGSLRTERLKITRMHTSDNQKPYVSRCSROCOBQ 507
 DB 485 TPGRYEIMNLQYVEFGADYINVGSWHGOLSIDDYMQINRSD--WVLSVCSSEPCSKGE 542
 QY 508 VRRV-KGFHSCQYDVCDEAGSYRONPDDICTFCGODEWSPERSTRCFRRSRFLANGE 566
 DB 543 IKVIRKGGVSCWICTACKDNEIYO--DEFTCTACDGLGMPDDPELEGEPTTLRLLEWGN 600
 QY 567 PAVLLILLLSLAELVLAALGLFVHNHSDPLVQASGGLACFGL---VCLGLVLCVILL 623
 DB 601 PESIVQVFAELGLITVTSFVTEIFVLYRDPYVASSREL-CYTLAIFLGLYICPFLII 659
 QY 624 FPGQSPARCIAQOPLSHLPLTGLSTLFLQAEIIEVESELPLSWADRLSG-----CLRG 678
 DB 660 --AAPTVAASYCLORLLVGLSATMWCYSALVTKTNR-----ARILASGKKIKCTRK 707
 QY 679 P-----NAMLVVL-LAMLYEVALCTWYLVAFPRPEVYTDMMHLPTALVHCTRSMVSGL 732
 DB 708 PRFSAMQVLIAGLVSVQLTLEVTLLILEPMPVKSPST-REVFLICMT-STVGVA 765
 QY 733 AAHTNATLAFCLFELTVRSQPCGYNARAGLTFAMLAFTTWVSFVLL--ANQOVL 790
 DB 766 PLGNGLLIMCTYYAFTRVNPANENAKYIAFTMYTCIIIMLAFVYIYIGSNKII-- 823
 QY 791 PAVONGALLICVLGILAAFLPRCYLLMRQPLNTPPEF 829
 DB 824 --TTSVSLSVTVLALGCMFSPKIYIILAKPERNVSFAF 860

RESULT 10

A49874

metabotropic glutamate receptor 7 - rat

Alternate names: metabotropic glutamate receptor mglur7

C:Species: Rattus norvegicus (Norway rat)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999

C:Accession: A49874; 157954

R:Okamoto, N.; Hori, S.; Akazawa, C.; Hayashi, Y.; Shigemoto, R.; Mizuno, N.; Nakanis

J. Biol. Chem. 269, 1231-1236, 1994

A:Title: Molecular characterization of a new metabotropic glutamate receptor mglur7 c

A:Reference number: A49874; MUID:94117433; PMID:8288585

A:Accession: A49874

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-915 <RES>

A:Cross-references: GB:D16817; NID:9458728; PIDN:BA04092.1; PID:9458729

R:Saugstad, J.A.; Kinzie, J.M.; Mulvihill, E.R.; Segerson, T.P.; Westbrook, G.L.

Mol. Pharmacol. 45, 367-372, 1994

A:Title: Cloning and expression of a new member of the L-2-amino-4-phosphonobutyric a

A:Reference number: 157954; MUID:94195260; PMID:8145723

A:Accession: 157954

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Cross-references: EMBL:U06832; NID:9459657; PIDN:AAA20655.1; PID:9459658

C:Genetics:

A:Gene: MGLUR7

C:Superfamily: metabotropic glutamate receptor 4

C:Keywords: neurotransmitter receptor

Query Match 13.4%; Score 604.5; DB 2; Length 915;
 Best Local Similarity 23.9%; Pred. No. 3.5e-39;
 Matches 212; Conservative 169; Mismatches 386; Indels 121; Gaps 32;

QY 6 VIGSLMALLHNGTAPCLSQLRMKGIVYVGGLEPGEAEAGLRSTRSPSSVCTRF 65
 DB 20 VLEVLTCLVLAARQEMKATAPHSIRIBEDYTLGGFLPYHAKGPSV-----FCGDIK 71
 QY 66 SSGILMALMAKMAVEEINKNSDLLPGRLGYDLFDTCSEPVAMKPSLMFLAKAGSRDI 125
 DB 72 RENGTHREAMLYALDQNSDPNLLPNTVIGARIDFTCSRDYVALLQESTFEVALIQKDT 131
 QY 126 A-AVCNTQYQYR-----RYLAVIGPSSSELMATGKFFSFLMPQVSYGASMEILLAR 177
 DB 132 SDVRC--TNGEPVVFKEKQVVGIGAGSSVSIVANILRLFLQIPQISTASTABELSD 189
 QY 178 EFPFSEFRVPSRVQLTAAELLQEGWVVAALGSDDEYQOGSLISFALAARQICIAH 236
 DB 190 RYDFPSKRVVPPDSQADAMDIYKALGMVYVSTLASGSGYGEKVESTQISKRAGLIC 249
 QY 237 IAHGELVPLPRADSRIGKQVDVLHQ-VNQGSSVQVLLFAS-VHAHALFNYSISRLSP 294
 DB 250 IAQSVRIQERKD--RTIDFRIIKQLDTPRSRAVIVFANEDIKQILAAKRAYQVGH 307
 QY 295 KVVVASEAWLT-----SDLVNG-LPGMAQMGYVLF-----LQKQGLHEFPQY 337
 DB 308 FLWVGSDSWGSKINPLHGHEDIAEGAITIQPRATVVEGFDAYTSRTLENNRNWFAEY 367
 QY 338 VKTHIALMDPAFCALGEROGLLEDVVGQRCPCDCITLQNVASGLNHHTFS--VY 394
 DB 368 WEENRCK-----LTSSSKEDIDRKCTGQ-----ERIKDSNYDEGVQYVI 412
 QY 395 AAVYSVAQALHN--TLQCNASG--CPAADPYKPMQLLENMTNLTPH-VGGLPLRDESSGN 449
 DB 413 DAVYAMAMALHNMKNDLQADRYGVCPEMEQAGKKLLKTYIRIVNNGSAGPYVMENKNGD 472
 QY 450 VDMETDLKLMWQGSV-PRLDVGRFNGSLRTERLKITRMHTSDNQKPYVSRCSROCOBQ 508
 DB 473 AGRYDIPOYQYTTNTTNGYRLIGOWTDELQINIEDMOMGVREIPSSVCTLLPCKPQR 532
 QY 509 RRVGFHSCQYDVCDEAGSYRONPDDICTFCGODEWSPERSTRCFRRSRFLAWGBPA 568

```

Db 533 KTKQKGPCCWTCPCD--GYOYOFDEMTCHCHCPYDOPNENRNGCCNIPITIKLEHSPW 590
Qy 569 VILLLLLLSTALGLVLAALGLFVHHRDSPLYOASGGLAGLGLVGLVGLVGLVGLVGLV 628
Db 591 AVIVFLAMGIATFIVMATFIRNDPIYRASGRELSYVLLGLVGLVGLVGLVGLVGLV 650
Qy 629 SPARCLAOOPSLHPLPGCLSTLEL-----QAEIIVESELSPLSMADRLSGCLGPMWL 683
Db 651 DVAICSRRAVE--LGLMGCSYALLTKTNIRIYRIFEGKKSVT--APRLI-----SPISQL 703
Qy 684 VLLAMLYEVALC---TWLYAPPEVYTDH---MLPEALVHCRTKRWVSGLAHAT 736
Db 704 AITSSLSIVQLGVEITFEGVD--PPNIIIDYDEKRTNPEQAR-----GVLKCD 750
Qy 737 NATLAFCFCGLTFV-----RSQGCYNRAGLTFALAYFITWVSFVPL-LA 783
Db 751 ITDQITISLGLSTILMATCVIAIKTRGVPEENNEAKPIGFTWYTCIWLAFIPIFG 810
Qy 784 NVOYVLRPAVOMGALLCV-----LGIILAFHLPRCYLLMRPGIN 824
Db 811 TAGSAEKLYIQTTLTITSMNLASVALGML---YMPKYIITIFHELN 855

```

RESULT 11

```

JH0562
metabotropic glutamate receptor 3 precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 13-Sep-1998
C/Accession: JH0562
R/Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
Neuron 8, 169-179, 1992
A>Title: A family of metabotropic glutamate receptors.
A/Reference number: JH0561; M01D:92110002; PMID:1309649
A/Accession: JH0562
A/Molecule type: mRNA
A/Residues: 1-879 <TAN>
A/Experimental source: brain
C/Comment: This protein is coupled to a G protein and evokes a variety of functions by m
C/Superfamily: metabotropic glutamate receptor 4
C/Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-879/Product: metabotropic glutamate receptor 3 #status predicted <MET>
F:577-599/Domain: transmembrane #status predicted <TRI>
F:614-634/Domain: transmembrane #status predicted <TRI>
F:646-664/Domain: transmembrane #status predicted <TRI>
F:689-709/Domain: transmembrane #status predicted <TRI>
F:735-756/Domain: transmembrane #status predicted <TRI>
F:770-791/Domain: transmembrane #status predicted <TRI>
F:804-828/Domain: transmembrane #status predicted <TRI>
F:209,292,414,439/Binding site: carbohydrate (asn) (covalent) #status predicted
F:610,845/Binding site: phosphate (Ser) (covalent) #status predicted

```

```

Query Match 13.2%; Score 597.5; DB 2; Length 879;
Best Local Similarity 25.2%; Pred. No. 1.2e-38;
Matches 218; Conservative 142; Mismatches 380; Indels 125; Gaps 32;

```

```

Qy 25 LSQQLRMKGQVYLGGLPLGE---AAEAGLRSTRSPVCTFRSSNGLLMALAMKAV 80
Db 29 MRRRIKIEGLVGLGFLPINKGTGTECGRINEDR-----GIORLEAMLFAI 76
Qy 81 EINKKSDLLPGLRGLYDLPDTCSEPVVAMKPSIMPL-AKAGSDIAAY-CNYTQYQ-- 135
Db 77 DEIKKDYVLLPGVGLYHIDTCSDRTYALQSLFVRASLTKVDEAYMCPDGSYAIQ 136
Qy 136 --PVIVL-VIGPHSSSLAMTGRKFFFLMPOVSYGASMEKLSAREPPSPFRVPSDRV 192
Db 137 NIPLLIGVIGGSYSVSIQVANIIRLFQIPQISYASTSAKLSKSRDYFARTVPDPFY 196
Qy 193 QLTAAAELEQFGNNVAAALGSDDEYGRQGLSIFSAALAAAGCTIAHGLVPLPRADSR 252
Db 197 QAKMAAILRFNWTYVSYVASEDYGTEGLEAFEGEARLRNICTATRAEKYGRNIRKSY 256

```

```

Qy 253 LGRVQDYLHOVNOSSVOVLLFASVHAHALFNYSISSRLSPK-VWVASEAW-LTSDLV 310
Db 257 DSVIRELLQKPN---ARVVLFMRSDDSDRELI--AAARNVASFTVVASDGMGAEISVK 311
Qy 311 GLPQAMQGVYGLFQAGADLHERPOYVKT-----HLATIDPAF-----CSALG 355
Db 312 GSEHVAAGAITLELASH--PVRQDRYFQSLNPNYNN-----RNWFDEWQKFOGSLON 365
Qy 356 EREGLEDVEDVQRCPCDCITLQNVASGLNHQTFEVSVAAYVSAOALH---NTLOCA 412
Db 366 KRNH-----ROYCDHLAIDSNTYE---QESIMVYVNAVYAMAHALHKMORTLCPNT 415
Qy 413 SG-CPAADPYKPMQL-ENMYNLTFHVGLP-----LRPSSGNDMEYDLKLYWOG 463
Db 416 TKLQDAKIIDGKKLYKEYLKIINFAPFNPKGADSVKFEFGDGMGRYVFNILQOTG 475
Qy 464 SVPRLDHVGRRNSLNRERLKRHTSDNOKPVSRCSQOEGQVRRVKGHSCCYCD 523
Db 476 GKYSYLKVGHAETLSLDVDSIHW--SRNSVPISQSDPCAPNEMKMOGPDVCCWICIP 533
Qy 524 CEAGSYRQNPDIACFCGODEMSPERSTRCFRRSRFLAMGEPAVLLLLSTLALGLV 583
Db 534 CE--PYEYLVDEFQCMDCGCGQWPTADLSCYNLPEDYIMEDMAIGPYTIACGLFCT 591
Qy 584 LALGLFVHHRDSPLYOASGGLAGLGLVGLVGLVGLVGLVGLVGLVGLVGLVGLV 643
Db 592 CIVTVFIRKNNMPLVYASGRELCYILFGVSLSYCTFEFFIAPSPVICALRLRIGLGS 651
Qy 644 LTGLSTLFLQA---AEIIVESELSPLSMADRLSGCLNG-----PMWLVYLLAM-LV 691
Db 652 FAITYSALLTKTNIAIF-----DVKAKGAPRKPTISSVOYFGLGLV 698
Qy 692 EVALCTWYLAFFPEVYTDHMLPTE---ALVHCRTSRWSFGIAHATNMTLAFELGT 748
Db 699 QIVWVSWVLIETPG--TRRYLPEKRETVLTKCNVAD--SMSLSTLYDVVLITCYVA 755
Qy 749 FIVASQGCYNRAGLTFALAYFITWVSFVPLLANOVVLRPAVOMGALLC----- 801
Db 756 FKTRKCEPNEAKFIFTWYTCIWLAFIPIF---YVTSDDYRVQTTMCTISVLSG 811
Qy 802 --VGIILAFHLPRCYLLMRPGIN 824
Db 812 FVVLGICLFA---PRVHLYLFOPOKN 833

```

RESULT 12

```

149142
metabotropic glutamate receptor 8 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C/Accession: 149142
R/Duvolsin, R.M.; Zhang, C.; Ramonell, K.
J. Neurosci. 15, 3075-3083, 1995
A>Title: A novel metabotropic glutamate receptor expressed in the retina and olfactory
A/Reference number: 149142; M01D:9529344; PMID:7722646
A/Accession: 149142
A/Molecule type: mRNA
A/Residues: 1-908 <RSS>
A/Cross-references: EMBL:U17252; NID:g854728; PIDN:AAA68149.1; PID:g854729
C/Genetics:
A/Gene: mglur8
C/Superfamily: metabotropic glutamate receptor 4
C/Keywords: neurotransmitter receptor

```

```

Query Match 13.1%; Score 593.5; DB 2; Length 908;
Best Local Similarity 23.4%; Pred. No. 2.3e-38;
Matches 206; Conservative 157; Mismatches 368; Indels 149; Gaps 29;

```

```

Qy 26 SQQLRMKGQVYLGGLPLGEAEAGLRSTRSPVCTFRSSNGLLMALAMKAVEINN 85
Db 37 AHSIRLDGDIILGGLFVYHAKGENGV-----PCQDLKKEKGIHLEMLTAIIDQTNK 88

```


Db 811 GCVVIGGLFA---PKVHVFQPOKN 833

RESULT 14

metabotropic glutamate receptor mGluR5 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999

C:Accession: A42916

R:Abu, T.; Sugihara, H.; Nawa, H.; Shigemoto, R.; Mizuno, N.; Nakanishi, S.

J. Biol. Chem. 267, 13361-13368, 1992

A:Title: Molecular characterization of a novel metabotropic glutamate receptor mGluR5

A:Reference number: A42916; MUID:92317054; PMID:1320017

A:Accession: A42916

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1171 <AB>

A:Cross-references: GB:D10891; NID:g220813; PIDN:BAA01711.1; PID:d1002186; PID:g220814

A:Experimental source: brain

A:Note: sequence extracted from NCBI backbone (NCBIIN:107749, NCBIIP:107750)

C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 12.9%; Score 584; DB 2; Length 1171;

Best Local Similarity 24.3%; Pred. No. 1.9e-37;

Matches 218; Conservative 154; Mismatches 388; Indels 136; Gaps 37;

31 MKGYVLGGLFPLGEAEAGLRSTRPSSPYCTPRSSNGLLMALAMKAVEINNKSDLL 90

31 MPGLIIIGALFVSNHOPVD-KVHERKCGAVREY---GIORVAMLTHERINSDFLL 86

91 PGLRIGYDLFTGSEPVVAMKPSLMF-----AKAGSRDIAYCNYQYOPRYLAV 141

87 PNITLGEIRDSCHMSAVALEQSTIEFIRDSLISEEEGLVRCYDSSSFRSKRP-IVGV 145

142 IGPHSSSELAMVTGKFFSFELMPQVSYGASMEELISARETFEFPFRVPSDRQVLAAL 201

146 IGPHSSSVAVIQVNLQLENIPOIAVSATSMDSKTLFKYFMVRVPSDAQARAMDIV 205

202 QEFMNVAALSGDDEYRGGLSIFSAALAAAGICIAHSGVLPVLRADSRIGKQVDTLH 261

206 KRYMTVYSAVHTGNGESGEMAEFKDMSAKEGICIAHSYKI-YSNAGEQSFKLTKLR 264

262 QVNOSSVOVLLFASVHAHALFNYSISSRLSPKVVAVASEAMLTSDVGLPQMAQMTV 321

265 S-HLPKRVAVACFEGMTVRGLL---MAMR---RLGLAGEF-----LLGSDGMADRYD 312

322 LGFLORGA-----QLHEFPQVY-----KTHLALADPAFCALSGERQ----- 359

313 TDGYQREAVGGITTKLOSPPDKWFDYYLKLREPNTL---RNPWFQEFWQHRCRLEGE 369

360 GLEEDVVGORCPQDCITLQNVASAGLNHHTFS---VYAAVYSAVALHNTLQCNASG- 414

370 AQEHSKYNKTCN---SSILTR-----THNVODSKMGFVYNATYSMAVGLHNMQMSLCGY 421

415 ---CPADPVKQWOLENNYNTLTF-HVGGPLRFEDSSGNVDMEDYDLKLMWQSVPLRD 470

422 AGLCDAMKPIDGRKILSLIMKTNFTGSGDMLTFDENDDSGRYEIMNFKEGK-----D 476

471 -----VGRF-NGSLRTERLRKIRWHTSDNOKPVSR--CSRCCGQGVARRV-KGHSCCYD 520

477 YFDYINWGSWNGELKMDDEV-WSKKN---ILRSVSECEKQKQIKVIRKEVSCWT 532

521 CVDEAGSYRONPDICTFCGQDEWSPERSTRCFRRSRRLANGEPVALLLLLLSTATL 580

533 CTPEKENEYV--DEYTCACQGLSWPDDTLTGCDLIPVQLRKNGDPRPIAAVYFACGL 590

581 GLVLAALGLFVHNRDPLVQASGGLACFGL---VCLGLVLSVLLFEGQSPARCIAAQ 637

591 LATLFVYVIFITVDTPVKKSSREL-CYIILAGICLGLCTFCLD--AKRKQIYCYQR 647

638 ---PLSHPLTGLCTSTLELQAAELFVSEELPLSMADR--LSGCLRGPRAMLVLLAMLYE 692

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

Db 648 IGIGLSPMAYSALVTNRIARILAGSKKRICKTRKPRFMSACQALVIAFLICIQIGIT 707

693 VALCTWYLVAFPPVYVTDWMLPTEALVHCSTRMSVSEGLAHATNATLAFICFGLVLR 752

708 VAL-----FIMEPPDMDYPSIREVYLICNTNIGVYTPGLY--NGLLISCFYFAKTR 761

753 SQPCYNRARGLTFAMLAFTWVSFVPL--ANOVV-----LRPVMGALLICVL 803

762 NVPANFNKAIATFTWTCIIMLAFVPIYGSNYKIKITMCFVSLSATVALGCMF----- 817

804 GILAFHLPRXYLMRQGLMTPEF-----LGGSPDAGQNDGNGNCGK 850

818 -----VPKVIILAPERNVRSATFTTIVYRMHVGDKSSSAASSSLVNLAK 866

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

||| : : : : : |||

```

OY 471 ---VGR--NSLKERLKIWMHSDNOKPVR--GROOCEQVRRV-KGFHSCYDCY 522
Db 480 DYINGSMDELKMDDEEV--WSKSN---IIRVCSEPECKGQIKYIRKEGVSCMCTC 535
OY 523 DCEAGSYRONPDDIACTEFCGODENSPERSTRCFRRSRFLANGEPAYILLLLSTALGL 582
Db 536 PCKENEYVF--DEYCKRCKQLGSMPTDDLIGCDLIPQYLRMDDPEPIAIVFACIGLLA 593
OY 583 VLAALGLEVHHRRDSFLVQASGGPLACGL--VCLGLVCLSVLLEPGQSPARCLAQO-- 637
Db 594 TLEVVVWFLIYRDPVPVAKSSSREL--CYIIILAGICLGYLCTFCCLI--AKPQIICYLQRTG 650
OY 638 -PLSLPLUTGCLSTPLLOAAEIEFESELEPLSMADR--LSCCLNGPMWMLVYLLAMLEYVA 694
Db 651 IGLSPMSYSALVYTKTNRIARILGSKKKICTKPKRPMSCAOVLVIAFLILCIGLGIYA 710
OY 695 LCTWYLVAFPEVVYTDWMHLPTEALVHCRTFRSWSFGLAHATNATLALFCILGTFTVRSQ 754
Db 711 L---FINEPDDIMHDPYSINEVYLLICNTNMLGVVPELgV--NGLILSTFFAFKTRNV 764
OY 755 PGCYRRANGLFAMALAFITTVWSVPLL--ANOVV-----LRPAVOMGALLCYLGI 805
Db 765 PANFNEAAYIAFTYTCIILAVLAPVIFYSNSYKIIMCFSVSLASATVALGCMF----- 818
OY 806 LAALFLPCYLLMROPLNPEPF-----LGGPBGDAQONGNNTGNQOK 850
Db 819 -----VKRYVYIIILKPERNVSAFTTSTVVRMHAGDDKSSSAARSSSLVNLMK 867

```

Search completed: May 19, 2003, 09:51:18
Job time : 25.1657 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:46:50 ; Search time 11.0911 seconds
(without alignments)

3186.136 Million cell updates/sec

Title: US-09-927-315-15
Perfect score: 4524
Sequence: 1 MGPAYGLSLMLHRCGTC.....GPGDAQGQNDGNTGNGCKHE 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1068.5	23.6	1078	1 CASR_HUMAN	P41180 homo sapien
2	1062	23.5	1079	1 CASR_MOUSE	Q99y96 mus musculu
3	1060	23.4	1085	1 CASR_BOVIN	P35384 bos taurus
4	1054	23.3	1079	1 CASR_RAT	P48442 rattus norv
5	670.5	14.8	872	1 MGR2_RAT	P31421 rattus norv
6	658.5	14.6	872	1 MGR2_HUMAN	Q14416 homo sapien
7	631.5	14.0	871	1 MGR6_RAT	P35349 rattus norv
8	622	13.7	912	1 MGR4_HUMAN	Q14833 homo sapien
9	619.5	13.6	877	1 MGR6_HUMAN	P31423 rattus norv
10	617	13.6	915	1 MGR7_HUMAN	Q14831 homo sapien
11	612.5	13.5	915	1 MGR7_HUMAN	Q14831 homo sapien
12	612.5	13.5	915	1 MGR7_HUMAN	Q14831 homo sapien
13	604.5	13.4	877	1 MGR3_HUMAN	P31422 rattus norv
14	604.5	13.4	915	1 MGR3_HUMAN	P31422 rattus norv
15	597.5	13.2	879	1 MGR3_RAT	P31422 rattus norv
16	594.5	13.1	908	1 MGR8_RAT	P47743 mus musculu
17	593.5	13.1	908	1 MGR8_MOUSE	Q99630 caenorhadi
18	585.5	12.9	908	1 MGR8_HUMAN	Q99630 caenorhadi
19	584	12.9	1203	1 MGR5_RAT	P31424 rattus norv
20	580.5	12.8	1212	1 MGR5_HUMAN	P31424 rattus norv
21	574.5	12.7	1199	1 MGR1_RAT	P23385 rattus norv
22	572.5	12.7	1194	1 MGR1_HUMAN	Q13325 homo sapien
23	474	10.5	999	1 MGR1_MOUSE	Q09630 caenorhadi
24	214	4.7	961	1 GBR1_HUMAN	Q99630 caenorhadi
25	212	4.7	960	1 GBR1_MOUSE	Q99630 caenorhadi
26	198.5	4.4	991	1 GBR1_RAT	Q99630 caenorhadi
27	175	3.9	1103	1 CYGF_BOVIN	Q02740 bos taurus
28	173	3.8	940	1 GBR2_RAT	Q08887 rattus norv
29	167.5	3.7	941	1 GBR2_HUMAN	Q07589 homo sapien
30	159.5	3.5	1108	1 CYGF_HUMAN	P31841 homo sapien
31	151.5	3.3	1108	1 CYGF_RAT	P31841 homo sapien
32	135.5	3.0	1323	1 NME4_RAT	Q06645 rattus norv
33	130	2.9	938	1 NMZ1_HUMAN	Q05586 homo sapien

34	129	2.9	1323	1 NME4_MOUSE	Q03391 mus musculu
35	128	2.8	938	1 NMZ1_RAT	P35439 rattus norv
36	125.5	2.8	938	1 NMZ1_MOUSE	P35438 mus musculu
37	118.5	2.6	1099	1 CYA7_MOUSE	P51829 mus musculu
38	117.5	2.6	1206	1 VGLM_RVRY	P03518 rlf1 valley
39	116	2.6	1336	1 NME4_HUMAN	P03518 rlf1 valley
40	115	2.5	986	1 CYGR_ARBPV	P15399 homo sapien
41	112.5	2.5	1197	1 VGLM_RVRY	P11528 arabaca pun
42	109	2.4	1110	1 CYGD_BOVIN	P21401 rlf1 valley
43	108.5	2.4	814	1 ADI5_HUMAN	P55203 bos taurus
44	108.5	2.4	918	1 ADI5_HUMAN	Q13444 homo sapien
45	108	2.4	836	1 GLK1_MOUSE	P39086 homo sapien
					Q60934 mus musculu

ALIGNMENTS

RESULT 1
ID CASR_HUMAN STANDARD: PRT: 1078 AA.
AC P41180: Q13912; Q16379; Q16108; Q16109; Q16110;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Extracellular calcium-sensing receptor precursor (Casr) (Parathyroid
DE Cell calcium-sensing receptor).
GN CASR OR GPRC2A OR PCAR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pearce S.H.S., Thakker R.V.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA TISSUE-Parathyroid;
RX MEDLINE=95279439; PubMed=7759551;
RA Garrett J.E., Capuano I.V., Hammerland L.G., Hung B.C., Brown E.M.,
RT Hebert S.C., Nemeth E.F., Fuller F.;
RL "Molecular cloning and functional expression of human parathyroid
RL J. Biol. Chem. 270:12919-12925(1995).
RN [3]
RP SEQUENCE FROM N.A.
RA TISSUE-Kidney;
RX MEDLINE=95408281; PubMed=7677761;
RA Aida K., Koishi S., Tawata M., Onaya T.;
RL "Molecular cloning of a putative Ca(2+)-sensing receptor cDNA from
RL human kidney.";
RN Biochem. Biophys. Res. Commun. 214:524-529(1995).
RN [4]
RP SEQUENCE FROM N.A.
RA MEDLINE=96343808; PubMed=8756555;
RA Freichel M., Zink-Lorenz A., Holltisch A., Hafner M., Flockerzi V.,
RT Raue F.;
RL "Expression of a calcium-sensing receptor in a human medullary
RL thyroid carcinoma cell line and its contribution to calcitonin
RL secretion.";
RN Endocrinology 137:3842-3848(1996).
RN [5]
RP SEQUENCE OF 643-908 FROM N.A.
RA MEDLINE=96193893; PubMed=8613532;
RA Birkle D.D., Ratnam A., Mauro T., Harris J., Pillai S.;
RT "Changes in calcium responsiveness and handling during keratinocyte
RT differentiation. Potential role of the calcium receptor.";
RL J. Clin. Invest. 97:1085-1093(1996).
RN [6]
RP VARIANTS FHH GLN-185; LYS-297 AND TRP-795.
RX MEDLINE=94094324; PubMed=7916660;
RA Pollak M.R., Brown E.M., Chou Y.-H.W., Hebert S.C., Marx S.J.,
RA Steinmann B., Levi T., Seidman C.E., Seidman J.G.;

RT "Mutations in the human Ca(2+)-sensing receptor gene cause familial
RT hypocalcemic hypercalcaemia and neonatal severe
RT hyperparathyroidism." ;
RN Cell 75:1297-1303(1993).
RN [7]
RP VARIANT ADH ALA-127.
RX MEDLINE=95179179; PubMed=7874174;
RA Pollak M.R., Brown E.M., Estep H.L., McLaine P.N., Kitor O., Park J.,
RA Hebert S.C., Seldman C.E., Seldman J.G.,
RA "Autosomal dominant hypocalcaemia caused by a Ca(2+)-sensing receptor
RT gene mutation." ;
RT Nat. Genet. 8:303-307(1994).
RN [8]
RN VARIANTS FHH MET-62; CYS-66; MET-138; GLU-143 AND GLN-227.
RX MEDLINE=95243222; PubMed=7776161;
RA Chou Y.-H.W., Pollak M.R., Brandi M.L., Toss G., Arnyqvist H.,
RA Atkinson A.B., Papapoulos S.E., Marx S., Brown E.M., Seldman J.G.,
RA Seldman C.E. ;
RA "Mutations in the human Ca(2+)-sensing-receptor gene that cause
RT familial hypocalcemic hypercalcaemia." ;
RL Am. J. Hum. Genet. 56:1075-1079(1995).
RN [9]
RN SEQUENCE OF 1-61 FROM N.A., AND VARIANT FHH ALA-39.
RX MEDLINE=95403641; PubMed=7673400;
RA Aida K., Koishi S., Inoue M., Nakazato M., Tawata M., Onaya T. ;
RA "Familial hypocalcemic hypercalcaemia associated with mutation in the
RT human Ca(2+)-sensing receptor gene." ;
RL J. Clin. Endocrinol. Metab. 80:2594-2598(1995).
RN [10]
RN VARIANTS NSHPT LEU-227 AND TYR-598.
RX MEDLINE=96292293; PubMed=8675633;
RA Pearce S.H.S., Trump D., Wooding C., Besser G.M., Chew S.L.,
RA Grant D.B., Heath D.A., Hughes I.A., Paterson C.R., Whyte M.P.,
RA Thakker R.V. ;
RA "Calcium-sensing receptor mutations in familial benign hypercalcaemia
RT and neonatal hyperparathyroidism." ;
RT J. Clin. Invest. 96:2683-2692(1995).
RN [11]
RN VARIANTS ADHP THR-116; HIS-681 AND SER-806, AND VARIANT SER-851.
RX MEDLINE=96311554; PubMed=8733126;
RA Baton J., Whiner K.K., Yanovski J.A., Cunningham A.W., Laue L.,
RA Zimmerman D., Cutler G.B. Jr. ;
RA "Mutations in the Ca(2+)-sensing receptor gene cause autosomal
RT dominant and sporadic hypoparathyroidism." ;
RL Hum. Mol. Genet. 5:601-606(1996).
RN [12]
RN VARIANT FHH ARG-174.
RX MEDLINE=97442275; PubMed=9298824;
RA Ward B.K., Stuckey B.G.A., Gutteridge D.H., Laling N.G., Pullan P.T.,
RA Ratajczak T. ;
RA "A novel mutation (L174R) in the Ca2+-sensing receptor gene
RT associated with familial hypocalcemic hypercalcaemia." ;
RL Hum. Mutat. 10:233-235(1997).
RN [13]
RN VARIANT FHH GLU-557.
RX MEDLINE=91603857; PubMed=11762699;
RA Nakayama T., Minato M., Nakagawa M., Soma M., Tohe H., Aoi N.,
RA Kosuge K., Sato M., Ozawa Y., Kamatsuse K., Kokubun S. ;
RA "A novel mutation in Ca2+-sensing receptor gene in familial
RT hypocalcemic hypercalcaemia." ;
RL Endocrine 15:277-282(2001).
RN [14]
RN -1- FUNCTION: SENSE CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM.
CC
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
CC
CC -1- TISSUE SPECIFICITY: FOUND IN KIDNEY, BUT NOT IN BRAIN, LUNG,
CC LIVER, HEART, SKELETAL MUSCLE, OR PLACENTA.
CC
CC -1- DISEASE: DEFECTS IN CASR ARE A CAUSE OF FAMILIAL HYPOCALCAEMIC
CC HYPERCALCAEMIA (FHH) AND NEONATAL SEVERE HYPERPARATHYROIDISM
CC (NSHPT). TWO INHERITED CONDITIONS CHARACTERIZED BY ALTERED CALCIUM

CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -2- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: S67307; AAB29171.1; -
 DR PIR: S40476; S40476.
 DR InterPro: IPR001828; ANF_receptor.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G-PROTEIN_RECP_F3_1; 1.
 DR PROSITE: PS00980; G-PROTEIN_RECP_F3_2; 1.
 DR PROSITE: PS00981; G-PROTEIN_RECP_F3_3; 1.
 DR PROSITE: PS02529; G-PROTEIN_RECP_F3_4; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 1085
 FT DOMAIN 20 613
 FT TRANSSEM 614 636
 FT TRANSSEM 637 650
 FT TRANSSEM 651 671
 FT TRANSSEM 672 682
 FT TRANSSEM 683 701
 FT TRANSSEM 702 725
 FT TRANSSEM 726 746
 FT TRANSSEM 747 770
 FT TRANSSEM 771 793
 FT TRANSSEM 794 806
 FT TRANSSEM 807 829
 FT TRANSSEM 830 837
 FT TRANSSEM 838 863
 FT TRANSSEM 864 1085
 FT DOMAIN 91 91
 FT CARBOHYD 131 131
 FT CARBOHYD 262 262
 FT CARBOHYD 288 288
 FT CARBOHYD 401 401
 FT CARBOHYD 447 447
 FT CARBOHYD 469 469
 FT CARBOHYD 489 489
 FT CARBOHYD 542 542
 FT CARBOHYD 595 595
 FT CARBOHYD 1085 595
 SQ SEQUENCE 1085 AA; 121170 MW; 5066D8C9CD13E47 CR664;
 Query Match 23.4%; Score 1060; DB 1; Length 1085;
 Best Local Similarity 30.0%; Pred. No. 3.5e-67;
 Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

DB 182 KSLRTIPNDHQATAMADIEYFRMMWGTIAADDDYGRGIGKFPREAREIDICIDS 241
 QY 241 GLVPLPRADSRICKVODVLEHNOSSVOVLLFASVAAHALNYSISSRSPKVVWAS 300
 DB 242 ELI-----SOYSEDEKIOQVVEIONSTAKYIVVSSGDPLEPLKEIYRRNITGRIMAS 297
 QY 301 EAMLTSLDYLKLPQMAQGVTLFQGAOLHEFPQVY-KHLLALNDPACSL----- 354
 DB 298 EAMSSSLIAMEYFHHVGTIGLAGQIPGFRFELQVHPKRSYHNGAKFEWET 357
 QY 355 -----GEREGL-----EEDVQGRCPQDCITLON 380
 DB 358 NCHIQEGAKGGLPDTFLRGHEEGARLSNPAPRLCTGEENISSVETPMYDTHLR 416
 QY 381 VSAGLNHHQFVSVAAYSVAAALHNTLQ-----NASCPADQVYKFWOLENNY 432
 DB 417 -----ISYNYLAVYSIAHALODITYCIPGRGLFTNGS-CADIKKVEAMOVYKILRH 467
 QY 433 LTFHVG-GLPLRFDSSGVNDEYDLKLM---VMQSVPLHDVGRFN-GSLRTERL----- 483
 DB 468 LNFYSNMGEOYTFDECGDLAENYSTIMNHLSPEDGSIV-FKEVGYINYAKKGERLIND 526
 QY 484 -KIRWHTSDNCKPVSRCSQOEGQVRR-VKGFSCCYDCVCEAGSYRONPDDIACFTC 541
 DB 527 EKILMSGFSREVPSPNSRDCIAGTRKGIIEGPTCECEVCEPDGEYSDETDASACDC 586
 QY 542 GQDESPSPSRTRCFRRNRFLKAGEPAVLLLLLSLALGLVLAALGLFYHHRSPLVQA 601
 DB 587 PDDFWSMNHNSCLAKELFELSWTEPFIATLRAVLGIFLTAVALGVIFRNTPIYKA 646
 QY 602 SGGPLA---CGVLQGLVCL-SVLLPFGQSPARLQAQOPLSHPLTGCTSTLELQAE 657
 DB 647 TNRELSTYLLFSLIC-----CFSSSLFEGEPQDTCRLRQAFGISPLCLISCLIVKINR 702
 QY 658 IFV--ESELPLS-----WADRLSGCLRGPMAMLVLLAMLVEVALCTWYLAAPPEVTD 710
 DB 703 VLVAFEAIPISFHRKWMGLNIQ-----FLVFLCTFMQIYCLAIMLTAPPSYRN 754
 QY 711 WHMLPTEAL-VHCRTSRVSGLAHATNATLAFCLFGLTFLVRSQPGCYNARAGITFAML 769
 DB 755 -HELEDELITFTCHBGSIMALGFLGYCLIAACFFPAFRSKRLEPFRNAKFTTESML 813
 QY 770 AYEFTWVSFVLLANVOVLLPRAVOMGALLICVGLIAFLAPRCYLLMRPGLTPE 827
 DB 814 IFFIWMISFIPAYASTYKFEVSAVEVLAIIAASGGLACIFENKVIITLFRSPNTIE 871
 RESULT 4
 CASR_RAT
 ID CASR_RAT STANDARD; PRT; 1079 AA.
 AC PA8442;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Extracellular calcium-sensing receptor precursor (CaSR) (Parathyroid
 DE Cell calcium-sensing receptor).
 OS CASR OR GPRC2A OR PCAR1.
 GN Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Kidney outer medulla;
 RX MEDLINE=95116508; PubMed=7816802;
 RA Riccardi D., Park J., Lee W., Gamba G., Brown E.M., Hebert S.C.;
 RT "Cloning and functional expression of a rat kidney extracellular
 RL calcium/polyvalent cation-sensing receptor.";
 RN Proc. Natl. Acad. Sci. U.S.A. 92:131-135(1995).
 RP
 RC SEQUENCE OF 1-294 FROM N.A.
 RC STRAIN=Wistar;
 RX MEDLINE=95241465; PubMed=7724534;

RA Ruat M., Snowman A.M., Snyder S.H.;
 RT "Calcium sensing receptor: molecular cloning in rat and localization
 to nerve terminals.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3161-3165(1995).
 CC -1- FUNCTION: SENSE CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
 CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
 G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U01034; AAC52149.1; -
 DR EMBL: U02089; AAC52195.1; -
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF000003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR002979; G_PROTEIN_RECEP_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECEP_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECEP_F3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_RECEP_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 1079 EXTRACELLULAR CALCIUM-SENSING RECEPTOR.
 FT DOMAIN 20 612 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 613 635 I (POTENTIAL).
 FT TRANSMEM 636 649 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 650 670 II (POTENTIAL).
 FT TRANSMEM 671 681 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 682 700 III (POTENTIAL).
 FT TRANSMEM 701 724 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 725 745 IV (POTENTIAL).
 FT TRANSMEM 746 769 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 770 792 V (POTENTIAL).
 FT TRANSMEM 793 805 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 806 828 VI (POTENTIAL).
 FT TRANSMEM 829 836 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 837 862 VII (POTENTIAL).
 FT TRANSMEM 863 1079 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1079 AA; 120867 MW; D76645503619736 CRC64;
 Query Match 23.3%; Score 1054; DB 1; Length 1079;
 Best Local Similarity 30.3%; Pred. No. 9; le-67;
 Matches 269; Conservative 163; Mismatches 363; Indels 92; Gaps 28;
 DB 9 LSIWALL-HFGTGAPLCLISQOLRMKGDYVGLFPLG---EAEAGLRSTRPSSPYCTR 64
 DB 8 LALLALAMHSSAGP---DRAOKKGDITLGGFLPIHFGVAADKDDKS--REPSVCIR 62
 QY 65 FSSNGLLMALAMKAVEINNKSDLPGLRIAGYDLFPTCSEPVYAMSLMFLA--KAGS 122
 DB 63 YNFRGFWLQAMIFAIKEINSSPLLPNMNLTGRIFPTCVTSKALEATLTSIFAQNKIDS 122
 QY 123 RDIATACNTQYQPRVLAVYGPISSELMATYTKFFSFFLMPQVSYGASMSLLSARETFPS 182

DB 123 LINDFEPCNSEHPISTIAVVGAGSGSVTAVANLGLFIYDQVYASSSRLSKNKYKS 182
 QY 183 FFRVPSPDRVQTLAAALLOEPGNWVAALGSDDEYRGOGISFSAALAAAGCITANEGL 242
 DB 183 FLRTIPDEHQATMAADITIEFRNNWGTIAADDYRGPIGIEFRFAEREDICIDSELS 242
 QY 243 VPLPRADSRIGKVQDYLHOVNOSSVOYVLLFASVBAHALFNYSISSRLSPKVVAVSEA 302
 DB 243 I-----SQYSDDEEIQVVEYIYONSTAKVIYVSSGPDLEPLIKIYARNITGRILASEA 298
 QY 303 WLISDLVGLPGMAQMGTVYGLQGAQLHEFPQYV-KTHIALATDPAFCSALGER--- 357
 DB 299 WASSSLTAMPEYFHVVGITGFGLKAGQIRGFRFLQKVPKRSVHNGFAKEFEETPNC 358
 QY 358 --EEGLE-----EDVVGORQCQD-----CITLQNYSA-----GLNHD-TE 391
 DB 359 HLDGAKGPLRPVDTFVASHDEGNRLNLSSTARPLCTGDBENINSVTPMDYEHRLISTY 418
 QY 392 SVYAAYVSVAAQALHNTLQC-----NASGCPADPEVKPMOLNNYNTLF-HVGLPL 442
 DB 419 NVYLAVYSIAHALQDIYTCLPGRGLFTNGS-CADIKVEAMOVYLKHLRHLNFTNMGEQV 477
 QY 443 RFDSSGNVDMEDYDLKLM---VMQSVPRLDYGRFN-GSLRTERL-----KIRNHTSDNQ 493
 DB 478 TFDECGDLVGNYSIINHLSPEDGSIV-FREVGYYNAYAKGERLFTNEKILMSGFSRE 536
 QY 494 KPYRSCROCOEGOVNR-VKGFHSCYDCYDCAGSYRONDDIACFFCODEMSPRST 552
 DB 537 VPSNCRDQOAGIRKRIITSEGPCCEYCBPGRGEISGETDASACCKCDDDFMSNHT 596
 QY 553 RCFFRRSRFLAMEBPVALLLLLLISALGLVLAALGLVHHRRSPLOVASGGPLA---CF 609
 DB 597 SCIAKEIEFLAWTEPFGIALTEFLAVLGIPTAVIAGVIFKFRMPYKATNRELSTYLLF 656
 QY 610 GYVCLGLVCL-VYLLEFGQSPARCQAQPLSHPLTGCLSTLELQAAELFV--ESLPL 666
 DB 657 SLIC---CESSSLEFFIGEPQDWTGRLRQPAFGISFVLCISCLIVKRNVLVEAKIPT 712
 QY 667 S-----NADRLSGCLRPMAMLVLLAMLEVALCTWYIVAFPEVYTDWMLPTEAL-V 720
 DB 713 SFHRKMGMLNQ-----FLVLECTPMOILCITWLTAPSSSRN-HELEDEITFI 763
 QY 721 HCTRISVSEGLAATNATLAFICLGTFLVRSQPGCYNRABGILTEFMAVFTWSEVP 780
 DB 764 TCHGSSILGSLIGYCLLALICLFFPAFRSKRPENPNAKFTTFMLIFFIWIISFIP 823
 QY 781 LLNVQVVLARPAVGMALLCYGIIAFLHPLRCYLLMRQPGINTPE 827
 DB 824 AYASTYGEVSAVEVIAIILASFGLLACIFPNKYITILFRPSRNTIE 870
 RESULT 5
 MGR2_RAT STANDARD; PRT; 872 AA.
 AC P31421;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, last sequence update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)
 DE Metabotropic glutamate receptor 2 precursor (mGluR2).
 GN GPM2 OR GPRC1B OR MGLUR2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA MEDLINE=92110002; PubMed=1309649;
 RX Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
 RL "A family of metabotropic glutamate receptors."
 Neuron 8:169-179(1992).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS

Db	166	ASTAKLSKSRDYEFARKTVPPDFQAKAMAEILRFNMWTVSVASSEGVDGTEIAFE	223
QY	228	ALAAAGCTCAHGEVLPRLPRADSRGKVDVLTQVNOSSVOYVLEFASVHAHALENYS	287
Db	226	LEAFARNICVATSEKVRAMSRRAFESEVYRALL--OKPSARAVALTRESEDAIRELT--A	280
QY	288	ISSRLSPK--VWVASEALVTSDIYVGLGGMOMGVLGFLQRGALHEFP--QYVKTILAL	344
Db	281	ATQRLNLSFWVWASDGGALLESVAGSERAAEGIT-----IELASYPLSDFSYFQSL	334
QY	345	-----ATDPAFCSALGEREGLEEDVVGORCPQDCITLQNVASGLNHHOTFSYAAVVS	399
Db	335	DPWNNSNNPF-----REWEEREFHCSFQRCOAANSLRAVPEQSKIMF--VYNAVYA	387
QY	400	VQAALNH---TLQCNASG--CPAADPVKPOLLEN-MTNLTFHWGGLP-----LRDSSG	448
Db	388	MAHLMHMRHALCPNTHTLCDAMPVNGRRLYKXKFLVNFEDAPFRPADTDDEVERPFRFG	447
QY	449	NVDMEDYLTLMWQGSVP--RLHDVGRNGSLRETRLKITNHT--SDNKKPVSRCSROQEG	506
Db	448	DGIGRIYETFLRAGSGRRYKQVGYAEGTLTDTSTIPWASPAQPLPMSRCEPCLQN	507
QY	507	QVRREKGFHSOCYDVCVCEAGSYRONDDIACITFCGODEMSPERSTRCFRRSRFLAMGE	566
Db	508	EKVSVOGEVCCWCITCQPYEYTL--DEFTCADCGGIGYPMNSLTGCFELPOEYIRMGD	565
QY	567	PAVILLILLSLALGVLVALGLFVHHRDSPLVOASGCPACISGLVCLVCL--SVLL	623
Db	566	AMAGPPTIACLGALATLFLVGLGFVRHNATPVVAKSGREL-CY--ILLGVLFLCYKMTFV	622
QY	624	FPGQSPARCLAQGPLSHLPPLTGLSTLFLQA---AEIF-----VESELPISMDRLSG	674
Db	623	FIAPSPSTAVCTLRRLGLTAFSVCSYSLKTRTNRIARIFGAREGARPRTISASQVAI	683
QY	675	CLRGPMMLVLLMLLEVALCTYVLAEP-----PEVTDHMHLPTEALVLCRT	724
Db	683	C-----LALISQILLIV--AMLVYEAQGTGETAPREREVYT-----LGCNH	723
QY	725	RSWVSFGLAHNTAATTAFLFCFLGTFLVRSQPCYNRARGLTAPMALAYITWVSFVPLLAN	784
Db	724	RDASMLC--SLAYVNLTLALCTLYAFKTRKCPENENAEKFTGFTWYTCIIMLAFLPIF--	780
QY	785	VQVVLRAVVOGMALLLC-----VLGILAAHLPRCYLIAMQPOLN-----TPPF	828
Db	781	--YTTSSDYRQTTMKCVSVSLSGSVVLCGLFA--PKLHILLFQPKNVVSHRAPYSRF	835
QY	829	FLGGPGDAGOGNDGTNGOG	849
Db	836	-----GSAAPRASANLIG-QG	849
RESULT 6			
MGR2_HUMAN			
ID	MGR2_HUMAN	STANDARD;	PRT; 872 AA.
AC	O14A16; Q9H3N6;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Metabotropic glutamate receptor 2 precursor (mGluR2).		
GN	GM2 OR GPRC1B OR MGLUR2.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_Taxid=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=9346007; Pubmed=7620613;		
RA	Flor P.J., Lindauer K., Puttner I., Ruegg D., Lukic S., Knopfel T.,		
RA	Kuhn R.;		
RT	"Molecular cloning, functional expression and pharmacological		
RT	characterization of the human metabotropic glutamate receptor type		
TL	2."		
TL	Eur. J. Neurosci. 7:622-629(1995).		

[2]
 RN SEQUENCE FROM N.A.
 RA Yasuyuki F., Akiko J.;
 RT "Structure and polymorphisms of the human metabotropic glutamate receptor type 2 (mglur2) gene: analysis of association with schizophrenia.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE ACTIVITY.
 CC MAY MEDIANE SUPPRESSION OF NEUROTRANSMISSION OR MAY BE INVOLVED IN SYNAPTOGENESIS OR SYNAPTIC STABILIZATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED IN DIFFERENT REGIONS OF THE ADULT BRAIN AS WELL AS IN FETAL BRAIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS. STRONGEST, TO MGLUR3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib-sib.ch).
 CC -----
 CC EMBL; L35318; AAA7685.1; -
 CC EMBL; AB045011; BAB19817.1; -
 CC Genew; HGNC:4594; GRM2.
 CC MIM; 604099; -
 CC InterPro; IPR001828; ANF_receptor.
 CC InterPro; IPR000337; GPCR_Mgr.
 CC Pfam; PF00003; 7tm_3; 1.
 CC PRINTS; PR00248; GPCR_MGR.
 CC PROSITE; PS00979; G_PROTEIN_REC_P3_1; 1.
 CC PROSITE; PS00980; G_PROTEIN_REC_P3_2; 1.
 CC PROSITE; PS00981; G_PROTEIN_REC_P3_3; 1.
 CC PROSITE; PS02559; G_PROTEIN_REC_P3_4; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 CC MultiGene family.
 CC SIGNAL 1 18 POTENTIAL.
 CC CHAIN 19 872 METABOTROPIC GLUTAMATE RECEPTOR 2.
 CC DOMAIN 19 567 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 568 590 I (POTENTIAL).
 CC DOMAIN 591 604 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 605 625 II (POTENTIAL).
 CC DOMAIN 626 636 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 637 655 III (POTENTIAL).
 CC DOMAIN 656 679 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 680 700 IV (POTENTIAL).
 CC DOMAIN 701 725 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 726 747 V (POTENTIAL).
 CC DOMAIN 748 760 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 761 783 VI (POTENTIAL).
 CC DOMAIN 784 793 EXTRACELLULAR (POTENTIAL).
 CC TRANSSEM 794 819 VII (POTENTIAL).
 CC DOMAIN 820 872 CYTOPLASMIC (POTENTIAL).
 CC TRANSSEM 820 872 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 547 547 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CONFLICT 12 12 P -> L (IN REF. 2).
 CC CONFLICT 210 210 A -> V (IN REF. 2).
 CC CONFLICT 496 496 A -> P (IN REF. 2).
 CC CONFLICT 748 748 N -> K (IN REF. 2).
 CC CONFLICT 776 776 L -> F (IN REF. 2).
 CC SEQUENCE 872 AA; 95507 MW; 058608C3C701EBD CXC64;
 Query Match 14.6%; Score 658.5; DB 1; Length 872;
 Best Local Similarity 27.4%; Pred. No. 6.3e-39;
 Matches 252; Conservative 131; Mismatches 381; Indels 155; Gaps 37;

QY 9 LSLMALHPGTGAPCLISQOLRMKGIVYGLIFPL----GEAEAGLRSRTSPSSVPCR 64
 DB 11 LPLMGAVAEGRPA-----KKVLTLEEDLVILGIFPVHQGGAEDCGPVEHNR----- 57
 QY 65 FSSKGLMALMKMAVEEINKKSDLLPRLRGIDYDFDPCSEPVYAMKSLFLKA----- 120
 DB 58 ---GIQRLKAMLEFALDRINDPHLLPGVRLGAHTLSDCSKDTNALKEQALDFVRSLSRG 113
 QY 121 --GSRDI---AAVCNYTOYOPRVLAIVIPHSSELAATMGKFFSPFIMQVSGASMETLS 175
 DB 114 ADGSHNICPDGSYARHGAPPAITGVIGGSYSIVQANILRLFOITQIYASISATLS 173
 QY 176 ARETFPSEFPYPSDRVQVLTAAELLOEFGNMYAALGSDDEYGRGSLISALAAAGI 235
 DB 174 DKSRDYDFARVPPDFQAKMAEILRFNMTYVTEASEDQGTGTEARELEARANI 233
 QY 236 CIAHGLVPLTRADDSRLGKQVDVLAHOVNOSSVOYVLLFASVHAHALFNTSISRLSPK 295
 DB 234 CVAATSEKGRAMSRRAEFGVYRALL--QKPSARAVLFTSEDAARELL--AASORLNAS 288
 QY 296 -VWVASEAM-LTSDLVMLPGMAQMGTVLFLQCAQLHEFPQYKTHAL----- 344
 DB 289 FTWVASDQMGALLESVYAGSEAAE-----GATIELASTPISDPASTYFQSLDPWN 338
 QY 345 -ATDPAFCSALGERQGLEEDVYQGRCP--QCDCT--TLQNVASGLNHQTFVYAAVYS 399
 DB 339 NSRNPWFREFEYQOR-----FRCSFRORDCAHSLRAVPEQESKIME--VYNAYVA 387
 QY 400 VAOALAHN---TLQCAVAG--CPADDPVKPMOLLEN--MYLFTFHVGLP-----LRFDSSG 448
 DB 388 MAHALHNHNRALCPNTYTLCDMPRVNGRLKDKDVLNWKDAPRPADTINEVAFDNRG 447
 QY 449 NVDMEDYDLKLVWQGVSP-RLHDYGRFNGSLRTERLKRIMHTSDNQKPY--SRCSROCE 505
 DB 448 DGIGYNIETYLKAGSGRYRQKGYMAEGTLTDSLTPW--ASPAGPLAASRCEPCLQ 506
 QY 506 GQVRVRFKFSCHDCDVCCEAGSTRQNDIACIFCGDEMSPEKSTCFRRSRFLAWG 565
 DB 507 NEVKSVOGVEYCCWVICPCQPEYERL--DEFTCADCGGAYPMNASTGCFELPOEYIRWG 564
 QY 566 EPAYVLLLLLSLALGLVLAALGLFVHRDSPVQASGPIACFLVGLVCL---SVL 622
 DB 565 DAMAVGPTIACGLALTLFVLGAVVRNATPVYKASREL--CY--ILLGVFLCYCTWF 621
 QY 623 LEPGQSPARLAQDPLSHPLTGCLSTFLQA--ABIF-----VESELPLSWADRLS 673
 DB 622 IFIAKPSAVCTLRRLGTFASVCYSALLKTRNIAIFGARGAGRAPRPFISPAQVA 681
 QY 674 GCLGPMALVYLLAMVEVALCTIYLAFP-----PEVYTDHMLPTALVHCR 723
 DB 682 ICL-----ALISQLLIYVA--WLVEAPGTGKETAPERREYVT-----LNCN 722
 QY 724 TRSWVFGIAHATNATLAFCLFGLFELVRSQPGCYNRARGLTFPMLAVFTWVSFVPLA 783
 DB 723 HRDMSMLD-SLAIVNVLALCTLTAFNTRKCPENNENKATIGFTYTTCTIWLALPLTF- 780
 QY 784 NVQVTLREAVQMGALLC-----VLGIIAFAHLPRCYLLAKROPGLN-----TPEFF 829
 DB 781 ---YTSSDYVHQVTTMCVSVLSGVSVLGCLFA---PKLIIILFQPKNVVSHRAPISR 834
 QY 830 LGGGPRGDAQGQNDGNTGQ 848
 DB 835 FGSAARASSSLGGSGSQ 853
 RESULT 7
 MGR6_RAT
 ID MGR6_RAT STANDARD; PRT; 871 AA.
 AC P35349;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 6 precursor (mglur6).

GN GRM6 OR GPCR1F OR MGLUR6.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Retina;
 RX MEDLINE=93280152; PubMed=8389366;
 RA Nakajima Y., Iwakabe H., Akazawa C., Nawa H., Shigemoto R.,
 RA Nakajima S.;
 RT "Molecular characterization of a novel retinal metabotropic glutamate
 RT receptor mglur6 with a high agonist selectivity for L-2-amino-4-
 RT phosphonobutyrate.";
 RL J. Biol. Chem. 268:11868-11873(1993).
 CC -1 FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 CC ACTIVITY.
 CC -1 SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1 TISSUE SPECIFICITY: RESTRICTEDLY EXPRESSED IN THE INNER NUCLEAR
 CC LAYER OF THE RETINA.
 CC -1 SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR4.
 CC -----
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 DR EMBL: D13963; BAA03066.1; -
 DR PIR: A46742; A46742.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE: PS0259; G_PROTEIN_RECP_F3_4; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Multigene family; Vision.
 FT SIGNAL 1 18
 FT CHAIN 1 871
 FT DOMAIN 19 871 METABOTROPIC GLUTAMATE RECEPTOR 6.
 FT TRANSMEM 19 579 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 580 602 I (POTENTIAL).
 FT TRANSMEM 603 616 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 617 637 II (POTENTIAL).
 FT TRANSMEM 638 648 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 649 667 III (POTENTIAL).
 FT TRANSMEM 668 691 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 692 712 IV (POTENTIAL).
 FT TRANSMEM 713 742 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 743 764 V (POTENTIAL).
 FT TRANSMEM 765 777 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 778 800 VI (POTENTIAL).
 FT DOMAIN 801 813 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 814 839 VII (POTENTIAL).
 FT DOMAIN 840 871 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 290 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 445 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 473 N-LINKED (GLCNAc. . .) (POTENTIAL).
 FT CARBOHYD 561 N-LINKED (GLCNAc. . .) (POTENTIAL).
 SQ SEQUENCE 871 AA; 95089 MW; 9E70B4D6A13B1867 CRC64;

Query Match 14.0%; Score 631.5; DB 1; Length 871;
 Best Local Similarity 25.6%; Pred. No 5.1e-37;
 Matches 225; Conservative 146; Mismatches 410; Indels 99; Gaps 31;

4 PAVIGLSMALHPGTGAPLCLSQLRMKGDYVIGLGPLGAEBAEAGLRSTRTPSSPYCT 63

DB 5 PVLILMLAMWLSQAGIA---CGAGSVRLAGLITGLGIPVHARGAAG-----RACG 52
 OY 64 RF-SSNGILLMALAMKMAVEEINKSDLLPGLRYGLYDPTCEPYVAMKPSLMF-----LA 118
 DB 53 ALKKEQGVHRLPAMLYALDRVNADELLPGVRLGARLLDPTCSRDYTALEQALSPFOALIR 112
 OY 119 KAGSDIAAYCYTYQYQY-----RYLAVTIPHSESLAMVTKKFPFLMQVSYGASM 171
 DB 113 GRGDDEDSVRCPGGVPLRSAPPERVAVAVASASVSIMAVNLRLEAIPOISASTA 172
 OY 172 ELLSAREFPSPFRTVPSPDRVOLGTAALDLEFGFNMVVAALGSDDEGROGSLFSALA- 230
 DB 173 PELSDSTRDFSRVPPDYQAQAMVDVRLAGNRYSTLASSEGNGSSGVAEAYQISR 232
 OY 231 AARGICIAHEGLVPLPRADSRIGVQDYLHOVNOVSQVQVLLPASYAAHALFNYSIS 289
 DB 233 EAGGVCIAS--IKIPR--EPKPGEFHKVIRRLMETPNARGLIIFANEDDIRRVLEATRO 288
 OY 290 SRLSPK-VWVASEAMULTS-DIVMGLPGMAQMGTVLGLQAGQLHEFPQYVTH----- 341
 DB 289 AMLTGHFLWVGSDSWGSKISPLINLEEA-VGALT-ILPKRASIDGFDQYFETRLNNR 346
 OY 342 ----LALATDPAF-----GSALGEREGLE-----EDVVGQRCPCQDCITLQNVASAGLNHO 389
 DB 347 RNWFAEFMEEFENFNCKINSSGQSDSTRKCGEERIGD-----SAVEQEK 394
 OY 390 TFSVTAAYVAQAALHNTLQCNASG---CPADPYKPMQLENNYNLTF-VGLPLRF 444
 DB 395 VQFVIDAVAYAIHALSHMQALCPGTGICPAMEPTDGTLLHYTRAVFNGSAGTPVWF 454
 OY 445 DSGNVDMEDYD-KLWVGQSVPR--LHDVGFPNSLIFERIKIRWHSNDOKVPSRCR 501
 DB 455 NENGAPGRITDFQYQATNGSSSGGYQAVGMAELRLDMEVYLRMGDPHPSPQSL 514
 OY 502 QCEQGVRR-VKGFHSCCYDCVDCSAGSYRQPDIDACTFCGODEWSPSTRCFRRSR 560
 DB 515 PCGPBERKKMVGV-PCMHCEACD--GYRFQVDEFTGACGDMRPPNHHGCAPTPV 571
 OY 561 FLAKBPAVLLLLLSLALGLVLAALGLFVHHRSPLVOASGGLACGLVLCGLVCS 620
 DB 572 RLWSSPMAALPLLAVALGIMATTTIMTFMRHNPTPIVIRASGRELSYVLLGIFLIVAI 631
 OY 621 VLLFGQSPARCLAOPLSHPLTNGCSTLELQA---AEIVSESLPMADRLSGCLR 677
 DB 632 TFLMAEPQALCAARLLILGTLTSLATLTNRIRYIEQGRKSTVTPPPIS---- 687
 OY 678 GPWMLVYL--LAMLVEVALCTWLVAPPEVYTDHMLPT-----EALVCRTRSWVS 729
 DB 688 -PTSQLVITFTGLTSLOVGVIAM-LGAQPPHSVIDYEEGRVDPQARGVLCMDSDSL 745
 OY 730 FGLAATATNTLFLFLGTLFLVRSQPCGNRARGLTFMALAFITVWVSFVPLANV---- 785
 DB 746 IGCL-GYSLILAMVTCTVYAIRKAGVPEFENKRGIFMTYTCIIMLAFVPIFGTAOSA 804
 OY 786 -QVLRPAVQMGALLCVGLIAAFHLPDRCYLLMBQPIN 824
 DB 805 EKITYIQTTLIVSLSASVSIGMLYPKTYVILFHPDON 844
 RESULT 8
 MGR4_HUMAN STANDARD: PRT; 912 AA.
 ID MGR4_HUMAN
 AC Q14833;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 4 precursor (mglur4).
 GN GRM4 OR GPCR1D OR MGLUR4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;


```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RA MEDLINE=96346635; PubMed=8738157;
RA Makoff A., Leichuk R., Oxer M., Harrington K., Emson P.;
RT "Molecular characterization and localization of human metabotropic
RL glutamate receptor type 4.";
RN Brain Res. Mol. Brain Res. 37:239-248(1996).
[2]
RP SEQUENCE FROM N.A.
RC MEDLINE=98141892; PubMed=9473604;
RA Mu S., Wright R.A., Rockey P.K., Burgett S.G., Arnold J.S.,
RA Roestek P.R., Jr., Johnson B.G., Schoepf D.D., Belagaje R.M.;
RT "Group III human metabotropic glutamate receptors 4, 7 and 8:
RT molecular cloning, functional expression, and comparison of
RT pharmacological properties in RGR cells.";
RL Brain Res. Mol. Brain Res. 53:88-97(1998).
[3]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RA MEDLINE=95342351; PubMed=7617140;
RA Flor P.J., Lukic S., Rueegg D., Leonhardt T., Knoefel T., Kuhn R.;
RT "Molecular cloning, functional expression and pharmacological
RT characterization of the human metabotropic glutamate receptor type
RT 4.";
RL Neuropharmacology 34:149-155(1995).
[4]
RN VARIANT ILE-797.
RP MEDLINE=21416233; PubMed=11525421;
RA Ohtsuki T., Toru M., Arihani T.;
RT "Mutation screening of the metabotropic glutamate receptor mGluR4
RT (GRM4) gene in patients with schizophrenia.";
RL Psychiatr. Genet. 11:79-83(2001).
CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
CC ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN THE CEREBELLUM.
CC EXPRESSED AT LOW LEVELS IN HIPPOCAMPUS, HYPOTHALAMUS AND
CC THALAMUS. NO EXPRESSION DETECTED IN LIVER.
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST TO MGURE.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
-----
DR EMBL; X80818; CAAS6784.1; -.
DR EMBL; U92457; AAB51762.1; -.
DR Genev; HGNC:4596; GRM4.
DR MIM: 604100; -.
DR InterPro; IPRO01828; ANF_receptor.
DR InterPro; IPRO00537; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PRINTS; PRO0248; GPCRMGR.
DR PROSITE; PS00979; G_PROTEIN_REC_P.F3.1; 1.
DR PROSITE; PS00980; G_PROTEIN_REC_P.F3.2; 1.
DR PROSITE; PS00981; G_PROTEIN_REC_P.F3.3; 1.
DR PROSITE; PS00982; G_PROTEIN_REC_P.F3.4; 1.
DR G-Protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Multigene family; Polymorphism.
FT SIGNAL 1 32
FT CHAIN 33 912 METABOTROPIC GLUTAMATE RECEPTOR 4.
FT DOMAIN 33 587 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 588 610 I (POTENTIAL).
FT DOMAIN 611 624 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 625 645 II (POTENTIAL).
FT DOMAIN 646 656 EXTRACELLULAR (POTENTIAL).

```

FT	TRANSMEM	657	675	III (POTENTIAL).	
FT	DOMAIN	676	699	CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM	700	720	IV (POTENTIAL).	
FT	DOMAIN	721	750	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	751	772	V (POTENTIAL).	
FT	DOMAIN	773	785	CYTOPLASMIC (POTENTIAL).	
FT	TRANSMEM	786	808	VI (POTENTIAL).	
FT	DOMAIN	809	821	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	822	847	VII (POTENTIAL).	
FT	DOMAIN	848	912	CYTOPLASMIC (POTENTIAL).	
FT	CARBOHYD	98	98	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	301	301	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	454	454	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	484	484	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	569	569	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	VARIANT	797	797	N->I.	
FT	SEQUENCE	912 AA;	101867 MW;	4A2F6E3A2EAF5A CRC64;	
SO	Query Match	13.7%;	Score 622;	DB 1; Length 912;	
	Best Local Similarity	23.8%;	Pred. No.2.5e-36;		
	Matches 224;	Conservative 139;	Mismatches 374;	Indels 204; Gaps 33;	
QY	4	PAYLGISLWALLHPGTGAPLC	-----	SOQLRNKGDYVGLGFRP 42	
DB	2	PKRRIGLGMW	-----	WAPRPLCLLSLYGFWPSSLGKRPGRHMSIRIDGITLGGFRP 56	
QY	43	LGAEADAGRSRTRRSPVCTRFSSNGLLMALAMKMAVEEINKRSDLLPGSLGVDLEPT		102	
DB	57	-----	VHGGSECKPCGGLKKEKGIHRLKEMALFALDRINDPDLNITLGRILDT	108	
QY	103	CSEPVYAKMPSIMLTAKAGSRD	---IAAYCN---	YTYQYPRVLAIVGPHSSLEAWTGFR 156	
DB	109	CSRDTHALEQSILTFQVALIEKDTFVRCGSGGPRITFRKERVGVIGAGSSSVSIMVANI		168	
QY	157	FEFLMPQVYSASMEILLSARETFPSFRFTVSDSDVOLTAAMELLEQFGMMVVALGSD		216	
DB	169	LRLEFIIPQISVASTAPDDNSRYPFESVWSDPYQAQAMVDYBALKMNIVSVASBG		228	
QY	217	EYGRGLSIFSLAALAAAR	---GICIAHEGLVLP	PRADDSRLGKVDVLDH--VMOSSVOVL 272	
DB	229	SYGEGGVEAF	---IQKSRREGGYCIANQ--	VKIPR--EPRAEGFDKIIIRLLETSSARAVI 282	
QY	273	LFAS	-----	VHAALFNSTISISRLSEPKVAVASEAMLTSDL	VMGLPRMAQMGTVLGF 324
DB	283	IFANEDDIRRVLEAAR	-----	AMQGHFEWMGSDSMGSKTAPVLRHLEEVAE	GAVT-I 334
QY	325	LORGALHEEPQVYTH	-----	LALATDPAFCSALGERDQGEEDVAVGRCPOC	373
DB	333	LPKRMSVGFEDYFESSRLTDNNRNIMWPAFEFEDNFHCL	---	SHALKKGSYHAKKCIJR 391	
QY	374	DCITLQNVASAGLNHQTESVYAAYVSVAAALHNTLQCNASG	---	CPAADPYKPMOLLEN 429	
DB	392	ERIGQD	---	SAYBQEKQVQFVIDAVYAMGHALHAMRDLCPRGVGLCPRMDDVDTQLLKY 449	
QY	430	MYNLTFH	---VGLRLAFEDSSGNDMEYDLKIMWQSSVPRLDHVGKFNGLRTERIKIMH 488		
DB	450	INNVFSGIAGPVTENEGDAPRDIYQYOLRNDSEAYKYIGSWTDHLRIERIMPR		509	
QY	489	TSDNKPYRSCROCOEGOVFR	---VGFHSCCDYCDVCEAGSYRQMPDDIACFFCGODEWS 547		
DB	510	GGGQQLPRLSICLPQRPGRKKTIVAG	---PCMCHCEPT---	GYQYVDYRTYCTCPCYDMRP 566	
QY	548	PERSTRCFRRRSRLTAKGEPAVLLLLLSLALGLVLAALGLFVHHRDSPLVQASGRLA		607	
DB	567	TNNRFGCRPIPIKILEMGSPMAVLLEFLFLAVVGIATFLFVILFVYVYNDPRTYKASG	---	622	
QY	608	CGIACVGLGVCLSVLLFPGQSPANCIAQOPLSHIPLTG	---	CLSTFLQAAE	657
DB	623	-----	RELSTVLAGIPLCYATTFPLMAEPLGCS	653	
QY	658	---IFVESLPLSMADRLSGCLR	---	GPMAVLVLLAM--LVEVA 694	

DB 654 LRRFLGLGMSISYALLKTKNRIRYIFEOGKRSVAPRPFISPAQSLATFSLISQLLQ 713
 QY 695 LCTWIVAPPEPVYDMMHLPTEALVHCRTSRVSGLAHAINATATFCFTGLV---- 751
 DB 714 ICWVEVVD-PSHSVDEQDQRTLDPRFAR-----GVLKCDISDLICLLGLYSMLNV 765
 QY 752 -----RSOPGCYNRARGLTFAMLAFTWVSFVPL-----ANYO 786
 DB 766 TCTVVAIKRTGVPEFENAKPIGFTMYTCIWMIAFIPFFGSGSADKLTYQTTLTVS 825
 QY 787 VVLRPAVQMGALLCVLGLTAAFLPRCYLLMRQPLNTP 827
 DB 826 VSLASVSLGML-----YMPKYIILFHPQNPVK 855
 RESULT 9
 MGR4_RAT
 ID MGR4_RAT STANDARD: PRT; 912 AA.
 AC P31423;
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 4 precursor (mglur4).
 GN GRM4 OR GPRC1D OR MGLUR4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92110002; PubMed=1309649;
 RA Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
 RT "A family of metabotropic glutamate receptors";
 RL Neuron 8:169-179(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93332699; PubMed=8338667;
 RA O'Hara P.J., Sheppard P.O., Thøgersen H., Venezia D., Haldeman B.A.,
 McGane V., Houamed K.M., Thomsen C., Gilbert T.L., Mulvihill E.R.;
 RT "The ligand-binding domain in metabotropic glutamate receptors is
 related to bacterial periplasmic binding proteins";
 RL Neuron 11:41-52(1993).
 CC -1 FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLYLATE CYCLASE
 ACTIVITY.
 CC -1 SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1 TISSUE SPECIFICITY: IS WIDELY DISTRIBUTED IN THE CNS. PREDOMINANT
 EXPRESSION IS SEEN IN THE GRANULE CELLS OF THE CEREBELLUM.
 CC -1 SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR6.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M92077; -; NOT ANNOTATED_CDS.
 DR EMBL: M90518; AAA93190.1; -;
 DR PIR: JH0563; JH0563.
 DR InterPro: IPR001828; ANF_receptor.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPRCMGR.
 DR PROSITE: PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE: PS00259; G_PROTEIN_RECP_F3_4; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KM Multigene family.
 FT SIGNAL 1 32
 FT CHAIN 33 912
 FT DOMAIN 33 587
 FT TRANSMEM 588 610
 FT DOMAIN 611 624
 FT TRANSMEM 625 645
 FT DOMAIN 646 656
 FT TRANSMEM 657 675
 FT DOMAIN 676 699
 FT TRANSMEM 700 720
 FT DOMAIN 721 750
 FT TRANSMEM 751 772
 FT DOMAIN 773 785
 FT TRANSMEM 786 808
 FT DOMAIN 809 821
 FT TRANSMEM 822 847
 FT DOMAIN 848 912
 FT CARBOHYD 98 98
 FT CARBOHYD 301 301
 FT CARBOHYD 454 454
 FT CARBOHYD 484 484
 FT CARBOHYD 569 569
 FT CONFLICT 124 124
 SQ SEQUENCE 912 AA; 101818 MW; 336430E19B4B577 CRC64;
 Query Match 13.7%; Score 619.5; DB 1; Length 912;
 Best Local Similarity 24.6%; Pred. No. 3; 8e-36;
 Matches 222; Conservative 145; Mismatches 397; Indels 137; Gaps 31;
 QY 4 PAVLGISLALLHFGT-----GAPICLSOOLRMKGIVYGLFPL-GEAEAGLRRTTRP 57
 DB 15 PCLLISLIVAPWVSSIGKPKGPHNNS--IRIDGITTGLGPVGRKSE----- 63
 QY 58 SSVPTCF--SSNGILMALMKAVEEINKNSDLPGLRLGYDLFCSEPVYAMKPSLMF 116
 DB 64 -GKAGEELKEKKGIIHLEMLFALDRINDPDLPIITIGARLIDCSNDTALBESLTF 122
 QY 117 LAKAGSRD-IAAYCN-----YTOYQPRVLAIVGPHSSSLAMVTKFFSFLMPQVSYGAS 170
 DB 123 VALLLEKDDTEVCGSGSPITTKPERVVGIVGASGVSSIVMANILRIKFIPOISYAST 182
 QY 171 MELLASRETFPSFRTVPDROVLTAAAELEFGNVAALGSDDEYGGOGSIFPSALA 230
 DB 183 APDLSDNSRITDFFSRVPSDITTOAQAMVDIVRLAKNNVSTLASSESYSGVEAFIQNS 242
 QY 231 AAR-GICIAHEGLVPLPRADDSRLGKRVODVLAH-VNOSVOVVLFLAS-----VHAH 281
 DB 243 RENGGVCIQNS--VKIPR--EPTGTGFDIRKLRLTSSNARGIILIPANEDDIRVLEAAR 298
 QY 282 ALFNISISRLSPKVVAVSEAM-LTSDLVMLGPMQMGTVLGEFLORGALHFEPOYVKT 340
 DB 299 R-----ANOTGHFFWMSDSMSKSAVPLRLVEVAE--GAVT-ILPKRMSVRFDFSS 350
 QY 341 H-----LALTDFAFGALGERGQLEEDVVGRCPCDDCITLQNVAGLNHHQ 389
 DB 351 RTLDNRRNIWPAEFEDNEHCL--SRHALKSGHIKCNIRERIGD--SAYQDEK 405
 QY 390 TFSVYAAVSVQAOLNHTLOCNAGS-----CPAADPKPMQOLEMNTLTFH--VGGLPLRF 444
 DB 406 VQFVIDAVYAMGHALHAMRDLCPRGVGLCPRPDPDVGQOLKIYNVNFSGIANGPVTF 465
 QY 445 DSSGNVMEYDLKLWVQGSVPRLDHVGPRNSLRTERLKIWNHSDNOKPVYSGRCGQO 504
 DB 466 NENGDAPEGRIYDYOYLNRGSAEYKIVGSWTDLHLIRIRMQPGSGOOLPFRSICLPQ 525
 QY 505 EGVYRR-VKGFHSCCYDVCDEAGSYRONPDDIACFFCODEMSPERSYRCFRRRRLFLA 563
 DB 526 PGRKRTVAGM-ACWNCBCT--GYOYVDRYTCCTCPYDMPTNRSQDPIPVKLE 582
 QY 564 WGPFAVLLLLLLSLALGLVLAALGLFVHHRDSPVLAQASGGPLACGLVCLGLVCLSVLL 623

Db 583 WDSFMAVLPFLAVVIGIAATLVVVFVRYNDPIKASGRSLSYLLAGIICVATFPL 642
 QY 624 FPGQSPARCLAQOPLSHLPGLSTFLQAAEIF-----VESELPWADRLSG 674
 Db 643 MIAEPDGTGSLRIEFLGSLGSIYALLKTKNTIYRLFEGKRSVAPRISASQLA 701
 QY 675 CLRGPMALVLLAMLEVALCTWYLVAPPEVVDNMLPTEALVHCRTSRVSVFGLAH 734
 Db 702 -----IFELISLQLLGICVWFVD-PSHSVDFODQRTLDPRAR-----GVLK 745
 QY 735 ATNATLAFCEIGETFLV-----RSQPCYNRABLTFLAMLAFTTWSPFPL 782
 Db 746 CDISDLSLCLGLGSLMLMTCTVYAKTRGVPEFENAKIGFTWTTCTVLAFLPIF 805
 QY 783 -----ANQVYLRAVOMGALLCVALGILAFHLPCCYLLMRQGLNTP 826
 Db 806 FGTSGASDKLYIQTTTLTVSVSLASVSLGML-----YMPKVIITLFPEDQNP 854
 QY 827 E 827
 Db 855 K 855

RESULT 10
 MGR6_HUMAN STANDARD; PRT; 877 AA.
 AC 015303;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 6 precursor (mglur6).
 GN GRM6 OR GPRC1F OR MGLUR6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97358610; PubMed=9215706;
 RA Hashimoto T., Inazawa J., Okamoto N., Tagawa Y., Beashe Y., Honda Y.,
 RA Nakanishi S.;
 RT "The whole nucleotide sequence and chromosomal localization of the
 RT gene for human metabotropic glutamate receptor subtype 6.";
 RL Eur. J. Neurosci. 9:1226-1235(1997).
 CC - FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 CC IS MODIFIED BY A G-PROTEIN THAT INHIBITS ADENYLYLATE CYCLASE
 CC ACTIVITY.
 CC - SUBCELLULAR LOCATION: Integral membrane protein.
 CC - SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR4.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U82083; AAB82068.1; -
 CC Genew: HGNC:4598; GRM6.
 DR MIM: 604096; -
 DR InterPro: IPR001828; ANF_receptor.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_RECEPTOR_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECEPTOR_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECEPTOR_F3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_RECEPTOR_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Multigene family; Vision.

FT SIGNAL 1 24
 FT CHAIN 25 877
 FT DOMAIN 25 585
 FT TRANSMEM 586 608
 FT TRANSMEM 609 622
 FT TRANSMEM 623 643
 FT TRANSMEM 644 654
 FT TRANSMEM 655 673
 FT TRANSMEM 674 697
 FT TRANSMEM 698 718
 FT TRANSMEM 719 748
 FT TRANSMEM 749 770
 FT TRANSMEM 771 783
 FT TRANSMEM 784 806
 FT TRANSMEM 807 819
 FT TRANSMEM 820 845
 FT TRANSMEM 846 877
 FT CARBOHYD 296 296
 FT CARBOHYD 451 451
 FT CARBOHYD 479 479
 FT CARBOHYD 567 567
 SQ SEQUENCE 877 AA; 95436 MW; AECDF91E8DA5994F CRC64;
 Query Match 13.68; Score 617; DB 1; Length 877;
 Best Local Similarity 24.98; Pred. No. 5,4e-36;
 Matches 212; Conservative 151; Mismatches 398; Indels 92; Gaps 31;

QY 29 LRMKGVYVGLGFLPGLAEAGLSRTRPSSPYTRFSSNGLMALAKMAVEINRKS 88
 Db 33 VRLAGGLTGLGFLFVHARGAAG-----RACGPL--KKEGVNHEALMALDRVNDPE 84
 QY 89 LLPGLRLGYDLFTGCEPVAAMKPSLMFL-----AKGSRDIAVYC--NTQYOP----R 137
 Db 85 LLPGVRRGARLIDCSHDTVALEQALSFVVALJRGSDGEGVRCGGVPLRPAPER 144
 QY 138 VLAIVGPHSSLEAMVTKKFSFELMPQVSYGASMLLSARETFSPRTVSDRVOFLAA 197
 Db 145 VVAVVGASASVSTIMVAVNLRLFAIPQISTASTAPELSDSTRYDFSRVVPDSDYQAM 204
 QY 198 AELLQEGWNVVALGSDDEYRGQSLIFSALA--AANGICIAHGLVPLPADDSRLGKY 256
 Db 205 VDIYRALGWNVSTLASEGNNGEGSEVAFVQISAEAGVCIAGS--IKIRP--EPKGEF 260
 QY 257 QDVLHGVNQS--SVQVLLFASVHAALFNYSISSRLSPK--VWVASAM--LTSDLVGLP 313
 Db 261 SKVIRLMEPPNARGIILFIANEDDIRVELAARQANLTGHFLVSDSPGAKTSPISLE 320
 QY 314 GMAQMGTVLGLQAGDALHEPQYVKH-----LALADPARCSALGEBQGLE 363
 Db 321 DVA-VGAILT-ILPRASIDGFDYFETRSLENNRNIMFEEWENCKL--TSSGTQS 376
 QY 364 DVVGQRPOCDCLTQVNVASGLNHQTFSS-----YAAVYSVAQALHNTLQCNASG-- 414
 Db 377 DDSIRK-----TGEERIGDSYIEQBGKQVFIADVAYLAHAHLSHQALCIGHT 427
 QY 415 --CPADPVKRWQLENNYNTLTFH--VGLPLREDSSGNVDMEYDL--KLWVQGSVPR--L 468
 Db 428 GLCPAMEPTGGRMLQYIRAVRFNGSAGTFVMEFNENDDAGRYDIFQYATNGSASSG 487
 QY 469 HDVGRFNGSLRTERLKIIRMTSDNQKPVSCRSQCGQVRR--VKGHSQCCYDQVDEAG 527
 Db 488 QAVGQMAETLRLDVALQWGSDDPHEVSSLSLPCGGEERKMKVKG--PCWCBEACD-- 544
 QY 528 SYRONPDIACTFCGQDEWSPERTSRCFRRRSRLFANGEPVALILLILLSLATGLVLAAL 587
 Db 545 GYRQVDEFTCEACPGDMRPPIPNHTGCRPIPVVALSSSPAAAPPLLAVALGIATTVV 604
 QY 588 GLFYHNRDSPLYQASGGLACFLVGLVCLVSLVLPFGQSPARCLAQOPLSHLPGLTC 647
 Db 605 ATFVRVNTPIVRSAGRELSTVLLTGIFLYAIFLVAWEAGVAVCARRLFGLGTTLS 664
 QY 648 LSTFLDA---AEIFVSELPISWADRLSGCLRPAMLVVL--LALILEVALCTWLVLA 702

Db 665 YSALLTKNRIRIIEFGKRSVTPPEFIS-----PTSQVITFSLISLVGMIAW-LGA 718
 QY 703 FPEPEVTOMHMLPT-----EALVHCRTSMWSFGLAHNTNLAFLCFLGFLVRSQSG 756
 Db 719 RPHSHVIDEBOORTVDPEQARGVLKCDMSDLISGL-CYSLLIMVTCIVAIKARGVE 777
 QY 757 CYNARGLTFAMLAFTVWSFVPLIANV-----QVLEPFAOMGLALLCYGIIAFLH 811
 Db 778 TENEKPIGFMTYTCIIMLAFVPIFFGIAQSAEKIYIOTITLVYSLISASVGLMYV 837
 QY 812 PRCYLLMRPGLN 824
 Db 838 PKTYVILFPEQN 850
 RESULT 11
 MGR7_HUMAN
 ID MGR7_HUMAN STANDARD: PRT; 915 AA.
 AC Q14831;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 7 precursor (mGluR7).
 GN GRM7 OR GPRC1G OR MGLUR7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96437220; PubMed=8840028;
 RA Makoff A., Pilling C., Harrington K., Emsen P.;
 RT "Human metabotropic glutamate receptor type 7: molecular cloning and
 RT mRNA distribution in the CNS.";
 RL Brain Res. Mol. Brain Res. 40:165-170(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96141892; PubMed=9473604;
 RA Wu S., Wright R.A., Rockey P.K., Burgett S.G., Arnold J.S.,
 RA Rostock P.R., Jr., Johnson B.G., Schoepp D.D., Belagaje R.M.;
 RT "Group III human metabotropic glutamate receptors 4, 7 and 8:
 RT pharmacological properties in RGT cells.";
 RL Brain Res. Mol. Brain Res. 53:88-97(1998).
 RN [3]
 RP VARIANT PHE-433.
 RX MEDLINE=21095249; PubMed=11163549;
 RA Bolonna A.A., Kerwin R.W., Munro J., Arranz M.J., Makoff A.J.;
 RT "Polymorphisms in the genes for mGluR types 7 and 8: association
 RT studies with schizophrenia.";
 RL Schizophr. Res. 47:99-103(2001).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 CC ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY AREAS OF THE BRAIN,
 CC ESPECIALLY IN THE CEREBRAL CORTEX, HIPPOCAMPUS, AND CEREBELLUM.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR4.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X94552; CA64245.1; -
 CC EMBL: U92458; AAB51763.1; -
 CC Genew: HGNC:4599; GRM7.
 DR MIM: 604101; -

DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF000003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_RECPEP_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECPEP_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECPEP_F3_3; 1.
 DR PROSITE: PS0259; G_PROTEIN_RECPEP_F3_4; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KM Multigene family; Olfaction; Polymorphism.
 KW SIGNAL.
 FT CHAIN 1 32
 FT DOMAIN 33 915
 FT TRANSMEM 591 615
 FT DOMAIN 616 627
 FT TRANSMEM 628 648
 FT DOMAIN 649 654
 FT TRANSMEM 655 675
 FT DOMAIN 676 702
 FT TRANSMEM 703 723
 FT DOMAIN 724 753
 FT TRANSMEM 754 775
 FT DOMAIN 776 810
 FT TRANSMEM 811 825
 FT DOMAIN 826 850
 FT TRANSMEM 851 915
 FT CARBOHYD 98 98
 FT CARBOHYD 458 458
 FT CARBOHYD 486 486
 FT CARBOHYD 572 572
 FT VARIANT 433 433
 SQ SEQUENCE 915 AA; 102250 MW; CFF94E06FF74919 CMC64;
 Query Match 13.6%; Score 615.5; DB 1; Length 915;
 Best Local Similarity 24.2%; Pred. No. 7.3e-36;
 Matches 215; Conservative 167; Mismatches 385; Indels 121; Gaps 32;
 QY 6 VLGSLMALHGTGAPLCLSOOLRKPKDYVIGGLFPLGLAEAGLRSTRSSPCTCF 65
 Db 20 VLEVLICALAAARQGEWAPHSIRIEGVTGLGFLPVARAKGSGV-----PGDIR 71
 QY 66 SSGGLMALAMKAVEINNKSDLPGLRGDLPCTGCEPVAMKPSIMFLAKAGSRDI 125
 Db 72 RENGIRLEAMLYALDQINSDPMLPNTGLARIDTCSRDYALQSLTFVQALIQKDT 131
 QY 126 A-AYCNYTQYOP-----RVLAIVGPHSEELAMVYGFSEFLMPQVSGAMELSAR 177
 Db 132 SDVRC--TNGEPFVFPKPKRYGVIGAGSSVSIMVANTLRFLQIDISYASTAPRLSD 189
 QY 178 ETPPSEFTVPSDRVOLTAABELDQFGNNWYAAALSGDDEYGRGQISITSAALA-AARGIC 236
 Db 190 RRYDFEFSRVPPDSFQAQAMVDIVKALGNVYSTLAESGSGYKVESTQISKEGAGLC 249
 QY 237 IAHGELVPLPADDSRLGAVQVOLIHO-VNOSQVYVLELAS-VHAHALFNYSISRLSP 294
 Db 250 IASQVRIPEKRD--RTIDFRIKQLDTPNSRAVYITANDEDIQTLLAAKRAQVGH 307
 QY 295 KYWVASEAMLT-----SDLVMG-LIPGMAQGVYIGF-----LQGAQLHEFPY 337
 Db 308 FLWVGSMSGSKINPLHGEDIAEGAITIQPRATVEGDAYFTSRLNNRNWVFAFY 367
 QY 338 VATHLALNDPAFCSLGRREGLEDVVGRCPCDCITLQNVASGLHHGTFSS---YV 394
 Db 368 WEENFNCK-----LITSGSKEDTDKCGO-----ERIGDSVYEDGKVQFYI 412
 QY 395 AAVYSVAQALHN--TLQCNASG-CPAQDPVPWQMLENNMYLTFH-VGGLPLRFPSSGN 449
 Db 413 DAVYMAHAHLHMKNDCLADYRGVCPDEMGAGSKILKIRIVNFGSGAGTYPMFKND 472
 QY 450 VDMEDYL-KLAWQGSVPLHDVGRNGSLRTERLIRNHTSDNCKPVYRCSROCOEGGV 508

Best Local Similarity 23.4%; Pred. No. 1.3e-35;
Matches 210; Conservative 159; Mismatches 380; Indels 147; Gaps 29;

QY 26 SQQLRMKGDYVGLGFLPGLAEAGLRSTRPSSPVCTRSSNGLLMALAMKMAVEEINN 85
D 36 SVSVSLPDDIILGFLPVEHKE-----GAPCGPKVYNRGQRLMALMLAIDRVNN 86
QY 86 KSDLLPGLRLGDLFDTCSEPVVAMKPSLML-AKAGSRDIAVCNTYQYQPR----- 137
D 87 DPNILPGITIGVHLDTCSDRPTALNOSIQFVRASLNNIDTSGVACADGSSFOLEKNASS 146
QY 138 --VLAVIPHSESLAMVGKFPFELMPOVSAGSMELLSAREPSPFRPVSPDRLT 195
D 147 GPVEVIGIGSVSVSLQVLANLRLPHIQVSPASTAKTSLDKTRDLARVPPDPTFSV 206
QY 196 AAELLOFGNMVVAALGSDDEYGRQGLSIFSAALAAAGICIAHGLVPLPRAD--DSR 252
D 207 ALVDILKFMNMSYVSTISEGSEYGEIHLKTEATERNVCIANAKEYPSAADKVPDSI 266
QY 253 LGKQVDVLAHQVNOSSVQVLLFASVHAHALEFNISIRLS-PRVWVASEAM-LTSDLYM 310
D 267 ISKLD-----KKPVAKGVVLEFTRAEDRRILQAKKRANLSQPFHMASDGGKQKLE 320
QY 311 GLPGMAQGTVLGFLORGAQLHEFPQY-----VKTHALA 345
D 321 GLEDIAE-GAITVELQ--SEIADPDRYMMQLPETNQNPFAEYMEDTFNCKVLSLVK 378
QY 346 TDPAFCSALGEREGLEDYVGQRCPCDCTITLONVSGNLHHQ--TFSYAANYSAQ 402
D 379 PDTSSANSTNKIGVK-----AKTECDYSRLSEKVGEDSKTQFVADYAYAY 430
QY 403 ALHN--TLQCN-----ASGCPADQPVKPMOLLEN-MY 431
D 431 ALHNHNRCHTQTSQTTETRKHLQSESVWTRKISTDTRKSOACPMANYDGEFYNNTL 490
QY 432 NLTF-HVGGPLREDSSGNVMEYDLKLMWQ--GSVPLHGVHFGNSLTERLKIWMH 488
D 491 NVSFDLAGSEVKFRQDGLARVDILNVQGENSSGVYKYGKWMFNGLOLNSTYVWN 550
QY 489 TSDNCKPVSRCQCEQVRRVKGFSGCCYDCVDCAGSTRONDDIACFTCGDEWSP 548
D 551 -KETEQPTSAQSLPCEVMIRKQOG-DTCWMCSCSESEFEVY--DEFTCKDCGGLMPY 606
QY 549 ERSTRCFRRSRFLAMGSPAVLILLLLSLALGLVLAALGLFVHHRDPIVOASGGLA- 607
D 607 ADKLSCTALDLYMKMNSLFLIPALIFGIALTSIVYVFAKNHDPVLRASGRELSTY 666
QY 608 --CFGLVCLGLVC-LSVLLFPGOPSPARCLAQO-----PLSHPLTGLCSTFLQ 654
D 667 TLIRGI-----LVYCNCPTALLAKPTIGSCVLRGIGVGSITYSALLKTNRISRTHS 722
QY 655 AAELFVESELPLSWADRLSGCLRGFWAMLVYLLAMLEVALCTWYLVAPPEVYTDWMML 714
D 723 ASKS-----AQRK--YISPOQVYVITSLDIAOVLTMTMVMVVEPRTREYYPD 770
QY 715 PTEALVHCRTSRWSVFGALHATNATLAFCLFGLTVSPGOCYRARGLFAMLAYFT 774
D 771 RREVLTKRKID-KSFLFSQLMMLITICITIAKTKRIENNESFIFGTMTTTCIT 829
QY 775 WVSFVPLLANYVLRPAVQMGALLCYLGILAA-----FHLPCYLLMRPGLIN 824
D 830 WLAEPVIFYFG-----GNSYEVOYTTLCISISLSASVALVCLSPRYVILVFPDKN 881

RESULT 13
MGR3_HUMAN
ID MGR3_HUMAN STANDARD; PRT; 877 AA.
AC Q14832;
DE 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metabotropic glutamate receptor 3 precursor (mglur3).
GN GRM3 OR GPRC1C OR MGLUR3.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE-96437205; PubMed-8840013;
RA Makoff A., Volpe F., Leichuk R., Harrington K., Emson P.;
RT "Molecular characterization and localization of human metabotropic
RT glutamate receptor type 3."
RL Brain Res. Mol. Brain Res. 40:55-63(1996).
CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
CC ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X77748; CA54796.1; -.
DR Genew; HGNC:4595; GRM3.
DR MIM; 601115; -.
DR InterPro; IPR001828; ANF_receptor.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PRINTS; PR00248; GPCRMR.
DR PROSITE; PS00979; G-PROTEIN_RECP_F3_1; 1.
DR PROSITE; PS00980; G-PROTEIN_RECP_F3_2; 1.
DR PROSITE; PS00981; G-PROTEIN_RECP_F3_3; 1.
DR PROSITE; PS0259; G-PROTEIN_RECP_F3_4; 1.
DR KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Multigene family.
FT SIGNAL 1 20
FT CHAIN 21 877
FT DOMAIN 21 574
FT TRANSMEM 575 597
FT DOMAIN 598 611
FT TRANSMEM 612 632
FT DOMAIN 633 643
FT TRANSMEM 644 662
FT DOMAIN 663 686
FT TRANSMEM 687 707
FT DOMAIN 708 732
FT TRANSMEM 733 754
FT DOMAIN 755 767
FT TRANSMEM 768 790
FT DOMAIN 791 800
FT TRANSMEM 801 826
FT DOMAIN 827 877
FT CARBOHYD 207 207
FT CARBOHYD 290 290
FT CARBOHYD 412 412
FT CARBOHYD 437 437
FT SEQUENCE 877 AA; 98619 MW; 66F28663CE35F740 CRC64;
Query Match 13.4%; Score 604.5; DB 1; Length 877;
Best Local Similarity 24.9%; Pred. No. 4.2e-35;
Matches 225; Conservative 151; Mismatches 372; Indels 157; Gaps 35;
QY 25 LSQQLRMKGDYVGLGFLPGLAEAGLRSTRPSSPVCTRSSNGLLMALAMKMAV 80
D 27 LRREIKIGDLYVGLGFLPINKGTGTECGRIENR-----GIORLEMLPAI 74
QY 81 EFINNKSDLLPGLRLGDLFDTCSEPVVAMKPSLML-AKAGSRDIAVCNTYQYQ-- 135

Db 75 DEINDDYLLGVKGLVHLLTSCSDTYALQSOSEFVRSALTKVDEAYMCPDGSYAIQ 134
 QY 136 --PRVLA-VIGPHSSELMTATGKFFSFLMPQVSYGASMEILLASRETPSEFRYPSDRV 192
 Db 135 NIPPLIACVIGSYSSVSIQVIANLLRLQIPIQIISASISAKISKRDIYFARYTPPPFY 194
 QY 193 QLTAAELLOEFGNMWNAALSDDEYSGROGISFASLAANGICIAHEGLVPLPRADDSR 252
 Db 195 QAKMAEILRFENMTYVSTVASEGDYGETIEAFQEARLNICIATAEKGRSNIKRSY 254
 QY 253 LGKVDVLYHOVNOSSVOYVLLF-----ASYHAHALFNYSISSRLSKVAVASEA 302
 Db 255 DSVIRELLQKFN--ARVVFPMNSDSRELLAASRNASF-----TWASDG 300
 QY 303 WLTSDLVGLPGMAQMGTVLGLQAGAO-LHEFPQVKT-----HLATDPAF----- 350
 Db 301 WGAQESIITKSEBHAVALGAIIT--LELASPVQFPDRIYPSLMPYNNH-----KNPWRDWE 354
 QY 351 ----CSALGEREOGLEEDYVQGRCPDCCITLQNVSAQINHQFES-----YAAVYVAQ 402
 Db 355 QKFOCSLQNKRNH-----RRV--CD-----KHLAIDSNYEORSKIMFYVNAVYAMAH 400
 QY 403 ALH--NTLQGNASG--CPADDPVKMOLLEN-MVNLTFHVGLP-----LRPSSGNV 450
 Db 401 ALHKMORTLCNNTTKLCDAKMLQOKKLYKLYLKINTAPPNPKDADSTVKEFTEDG 460
 QY 451 DMEYDLKIMWQGSYVRLHDYGRFNGSLRTERLKRMTSDNOKPVSRCQCOGQVRR 510
 Db 461 MGRNVNFQVWVGKYSYLAQGHMAETLSLVNSIHW--SNNSVTSQSCDPCAPAKENKN 518
 QY 511 VKGHSCCYDVCDEAGSYRKONPDDIACFCGQDEMSPERSTRCFRRSRLANGEPAYL 570
 Db 519 MQPBGVCCWICIPCE--PYEYLADEFTCMDCGSGQMPADLTGICYDLEDEYIRMEDAWAI 576
 QY 571 LLLLLLALGLVLAALGLFVHHSDPLVQASGGPLACFGVLCGLVSLVLPFGQSP 630
 Db 577 GPVITACIGFECTCMVAVVFTKHNNTPLVKASGRELCYITLFGVGLSCMTPEFLAKSP 636
 QY 631 ARCLAQOPLSHPLTGLCSTLFLQ--AEFVSESLPLSWADRLSGCLRG-----P 679
 Db 637 VICALRRGLGSSFAICYSALLTKTNCIARLF-----DYKNAQARKFISP 683
 QY 680 WAMLVVLLAM-LVEVALCTWLVAFPPRYVDMHMLPE--ALVHCRTSRWSVGLAHA 735
 Db 684 SSQVFLGLLVOIVMVSVMLEIAPG--TRRYTLAEKRETVILKICVWKO--SSMLISLT 740
 QY 736 TNATLAEFCIGTFLVRSOPGCYNRARGLTFAMLAFTWVSFVPLANVOVLRPANOM 795
 Db 741 YDVLIVLICVYAAKTRKCPENFEAKRTGTTMTTCIIMLAFPIF-----YVSSDYRV 796
 QY 796 GALLLC-----VIGIILAAFIHPRCYLLMRPGMLTPEF-----FLGGGPGDAQO 840
 Db 797 QTTMCISIVSLSGFVVLGLCPFA--PKVHILLPQOKNVVHRLHNFVSIGGTYSQ 853
 QY 841 MDGNT 845
 Db 854 SSAST 858
 RESULT 14
 MGR7_RAT STANDARD; PRT; 915 AA.
 AC P35400;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 7 precursor (mglur7).
 GN GRR7 OR GPRG1G OR MGLUR7.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;

RN [1] SEQUENCE FROM N.A.
 RP STRAIN-Sprague-Dawley; TISSUE-Brain;
 RC MEDLINE=94117433; PubMed=828685;
 RX Okamoto N., Horii S., Akazawa C., Hayashi Y., Shigemoto R.,
 RA Mizuno N., Nakanishi S.;
 RT "Molecular characterization of a new metabotropic glutamate receptor
 mglur7 coupled to inhibitory cyclic AMP signal transduction.";
 RL J. Biol. Chem. 269:1231-1236(1994).
 RP SEQUENCE FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE-Olfactory bulb;
 RX MEDLINE=94195260; PubMed=8145723;
 RA Saugstad J.A., Kinzie J.M., Mulvihill E.R., Segerson T.P.,
 RA Westbrook G.L.;
 RT "Cloning and expression of a new member of the L-2-amino-4-
 phosphonobutyric acid-sensitive class of metabotropic glutamate
 receptors.";
 RL Mol. Pharmacol. 45:367-372(1994).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: WIDELY DISTRIBUTED THROUGHOUT THE BRAIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR4.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D16817; BA004092.1; -;
 DR EMBL: U06832; AA20653.1; -;
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_REC_P3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_REC_P3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_REC_P3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_REC_P3_4; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Multigene family; Olfaction.
 FT SIGNAL 1 32
 FT CHAIN 33 915
 FT DOMAIN 33 590
 FT TRANSMEM 591 615
 FT DOMAIN 616 627
 FT TRANSMEM 628 648
 FT DOMAIN 649 654
 FT TRANSMEM 655 675
 FT DOMAIN 676 702
 FT TRANSMEM 703 723
 FT DOMAIN 724 753
 FT TRANSMEM 754 775
 FT DOMAIN 776 788
 FT TRANSMEM 789 810
 FT DOMAIN 811 825
 FT TRANSMEM 826 850
 FT DOMAIN 851 915
 FT CARBOHYD 98 98
 FT CARBOHYD 458 458
 FT CARBOHYD 486 486
 FT CARBOHYD 572 572
 SO SEQUENCE 915 AA; 102231 MW; F28AF4C6454AAC2 CRC64;
 Query Match 13.4%; Score 604.5; DB 1; Length 915;
 Best Local Similarity 23.9%; Pred. No. 4,4e-35;

Matches 212; Conservative 169; Mismatches 386; Indels 121; Gaps 32;

```

OY 6 VGLGSLMALHPGTGAPLGLSOLRKMDYVGLGFLGEAEAGLRSTRSSPCTCF 65
DB 20 VLEVLICVIAAARGEMAPHSIRIEGDTLGLFPPHAKGPBGV-----PGDIK 71
OY 66 SNGLLMALMKMAVEINNKSDLLFGLRGYDLFDTCEPVPVAMPSPIMFLAKASRDI 125
DB 72 RENGIRHLEAMLYALDQINSDPNLPLNYLGLARILDTCRDPYALAEQSITLQALIQKT 131
OY 126 A-AYCNYTOYQ-----RVLAIVGPHSELAMVTGKFFSEFLMPQVSYGASMELLSAR 177
DB 132 SDVRC--TNGEPPVFEVKPEKVVGVIGAGSSYSIMVANTILRLFOIPQISYASTAPLSD 189
OY 178 EFPSPFRVPSPDRQULTAAELLOFGNNVMAALGSDDEYRGQISTISALA--AARGIC 236
DB 190 KRIDFESRVPSPDFOQAMVDIVKALGNVYSTLASSESYGKVESTQISKEGGLC 249
OY 237 IAHGGLVPLPRADSRRLGKVQDVLHQ--VMOSSVQVLLFAS--VHAHALFNYSISRLSP 294
DB 250 IAGSVRIPEKRD--RTIDFRIIKQLDTPNSRAVIFRANEDIKQILAAKRAQVGH 307
OY 295 KVMVASEAMLT-----SDLVMG--LPGMAOMGYLGF-----LQGAQLHEFPQY 337
DB 308 FLWVSDSGSKINPLHGHEDIAEGAITIQPKRAVEGFDAFTSKTLENNRNWFAEY 367
OY 338 VTHIALANDPAFCALGREGLEDVYGRCPODDCTILQNVASGLNHQTFSS--YX 394
DB 366 WEENRCK-----LTISGKKEEDTRKCTGQ-----EIKGDSNYEGBKQVQYI 412
OY 395 AAVYSVAQALHN---TLQCNASG--CPAODPVKPMQLLENNMYLTFH--VGLPLRPDSGN 449
DB 413 DAVYAMAHLMHNMNDLADYRGVCEMEQAGKILKAKIRHYNENFSGTGYMFNKNDD 472
OY 450 VMEDDKLIMWQGSY--PLHYVGRNGSLRTERLKRHTSDNQPVRSRCSBOCEGY 508
DB 473 AGRGDIQYQTTNTNPGYRLIGOWTDELQNTIEMQMGKGVRELPSVCTLPCKPGOR 532
OY 509 RRAVGFHSCCYDCVCEASGYRONPDICTFCGODEMSPERSTRCFRRSRFLMGEPA 568
DB 533 KITQGTGTCWTCPCPD--GYOYQDEMTCCOCPYDQRENENRTGCONIPITKLEHSPW 590
OY 569 VLLLLLSIALGLVLAALGLFVHNHDSPLVQASGGPLACFGLVLCGLVSLLEPPGP 628
DB 591 AVIPEFLAMGLIATIFWATFIRNDPIVIRASGRELTVLLTGIFLCYITIFLMIAR 650
OY 629 SPARCLAQOPSLHPLTGLSTLFL-----QAAEIFVESELPLSWADRLSGCLRGWAWL 683
DB 651 DVAVCSFRFVF--LGIGMCISYAAALLTKNRIYRIFEGCKKSVT--APRLI---SPTSOL 703
OY 684 VVLLMLVAVVLC---TWLVAFPRPVYVDH---MLPTEALVHCRTSWVSFGLAHAT 736
DB 704 AL--TSSLISVQLLGVFIWGVDP--PNIITIDYDEHKTWNEQAR-----GVLKCD 750
OY 737 NATLAFLCFLGFLV-----RSQPCYNRARGILFAMLAIFYTWVSFVL--LA 783
DB 751 IYDLQIGSLGSIILMTCTVYAIKTRGVENENNAKIGFTMYTCTVLAIFIFRG 810
OY 784 NVQVILRAVQVQALLCY-----LGLLAFLPHLPCYLLMKRQPLN 824
DB 811 TAQSAEKLYITQTTLLISMNLASVATGML---YMPKVYIIFHPELN 855

```

RESULT 15

```

MGR3_RAT          STANDARD:      PRT:      879 AA.
ID  MGR3_RAT          P31422;
AC  01-JUL-1993 (Rel. 26, Created)
DT  01-JUL-1993 (Rel. 26, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Metabotropic glutamate receptor 3 precursor (mglur3).
GN  GRM3 OR GPRC1C OR MGLUR3.
OS  Rattus norvegicus (Rat).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=92110002; PubMed=1309649;
RX Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.:
RT "A family of metabotropic glutamate receptors."
RL Neuron 8:169-179(1992).
CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
CC ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: IS WIDELY DISTRIBUTED IN THE CNS. PREDOMINANT
CC EXPRESSION IS SEEN IN THE NEURONAL CELLS OF THE CEREBRAL CORTEX,
CC DENTATE GYRUS, AND GLIAL CELLS THROUGHOUT BRAIN REGIONS.
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
CC EMBL, M92076; -, NOT_ANNOTATED_CDS.
CC PIR, JH0562; JH0562.
CC InterPro: IPR001828; ANF_receptor.
CC InterPro: IPR000337; GPCR_Mgr.
CC Pfam, PF00003; 7tm_3; 1.
CC Pfam, PF01094; ANF_receptor; 1.
CC PRINTS: PR00248; GPCRMR.
CC PROSITE: PS00979; G_PROTEIN_RECPEP_F3_1; 1.
CC PROSITE: PS00980; G_PROTEIN_RECPEP_F3_2; 1.
CC PROSITE: PS00981; G_PROTEIN_RECPEP_F3_3; 1.
CC PROSITE: PS0259; G_PROTEIN_RECPEP_F3_4; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC Multigene family.
CC SIGNAL 1 22
CC FT CHAIN 23 879 METABOTROPIC GLUTAMATE RECEPTOR 3.
CC FT DOMAIN 23 576 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 577 599 I (POTENTIAL).
CC FT DOMAIN 600 613 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 614 634 II (POTENTIAL).
CC FT DOMAIN 635 645 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 646 664 III (POTENTIAL).
CC FT DOMAIN 665 688 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 689 709 IV (POTENTIAL).
CC FT DOMAIN 710 734 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 735 756 V (POTENTIAL).
CC FT DOMAIN 757 769 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 770 792 VI (POTENTIAL).
CC FT DOMAIN 793 802 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 803 828 VII (POTENTIAL).
CC FT DOMAIN 829 879 CYTOPLASMIC (POTENTIAL).
CC FT CARBOHYD 209 209 N-LINKED (GLCNC. . .) (POTENTIAL).
CC FT CARBOHYD 292 292 N-LINKED (GLCNC. . .) (POTENTIAL).
CC FT CARBOHYD 414 414 N-LINKED (GLCNC. . .) (POTENTIAL).
CC FT CARBOHYD 439 439 N-LINKED (GLCNC. . .) (POTENTIAL).
CC SEQUENCE 879 AA; 98959 MW; 3E5965DDE6DEED CRC64;

```

Query Match

Best Local Similarity 25.2%; Score 597.5; DB 1; Length 879;
Pred. No. 1.3e-34;

Matches 218; Conservative 142; Mismatches 380; Indels 125; Gaps 32;

```

OY 25 LSQGLMRKDYVIGLGFPGF-----AEAGLRSTRPSPVCRFSSNGLLMALAKMAY 80
DB 29 MREKITEGDLVYGLGFLPINEKGTGTECGRINEDR-----GIQRLAEMLFAI 76
OY 81 EEINNKSDLLPGRLGYDLFDCSEFPVAMKPSIMFL--AKAGSDIAAY--CNYTOY--- 135

```

Db 77 DEINKDNVLLPGVKGVLHLLDTCSDRFDYALRQSLFVRAASLTKYDEAEYMCPSGSAIOE 136
QY 136 --PRYLA-VIGPHSELAMVWGKFFSFILMPOVSGASMEILLARETFPSFFRTYPSDRV 192
Db 137 NIPLLIAGVIGSYSSVSIQVIANLRLFQIPQISTASTAKISDKSRDYFARTVPPDPY 196
QY 193 QLTAAELLQFEGNMVNAALGSDDEYGRQGLSIFSAALAAAGICIAHGLVPLPRADDSR 252
Db 197 OAKAAEILRFENNTYVSTVASEGDYGETEAFQEARLNINICATATAEKVGRSNIKRSY 256
QY 253 LGRVODVILHOVQSSVOYVLLFASVHAHALEFNYSISRLSPK-VWVASEAM-LTSDLYM 310
Db 257 DSVIRHLLQKRN--ARYVVLFRSDSDSRELI--AAANRVNASFTWASDGVGAQESIYK 311
QY 311 GLPGNAQGVTLGFLQGAOLHEFPQYKT-----HLALATDPAF-----CSALG 355
Db 312 GSEHVAIYGAITLELASH--PYRQFDRYFQSLNPNYNNH---RNPWFDRPWEQKFCQSLQ 365
QY 356 EREGLEDVVGORCPQDCITLQNVASAGLNHQTFSYAAYVSYAQAHL--NTLQCN 412
Db 366 KRNH-----ROYCDKHLAIDSSNYE---QESKIMFYVNAVYAMAHALHKMQRILCPNT 415
QY 413 SG-CPAODPVKFWOLL-ENMTNLTFHVGLP-----LREDSSGNVDMETDLKIMWOG 463
Db 416 TKLCDAMKILDGKLYKKEYLLKINFTAPFNPNKGADSIYKFTEDGDMGRYVFNMLQOTG 475
QY 464 SVPRLDHVRFGNSLRTERLKI RMTSDNOKPVSRCSRQCEGOVRAVKGFSHCCYCDVD 523
Db 476 GKYSIAKYGHAETISLDVDSIH--SRNSVPTSCSDPCAPENKKNMQPGDYCCWITCIP 533
QY 524 CEAGSYRONPDDIACFCGQDEWSPERSTRCFRRSRFLANGEPVALLILLLLSLALGY 583
Db 534 CE--PYEYLVDEFCTMDCGPGQFTADLSGCYNLEPDYIKMEDAWAIGPVTIACIGFLCT 591
QY 584 LAALGTFVHHDSPLVQASGGPLACFGVLCGLVCLSVILPFGQSPARCLAQBP LSHLP 643
Db 592 CIVITVFYKHNTPLVAKSGRELICYILFVSLSYCMTEFFLAKSPVICALRLRLGLGTS 651
QY 644 LTGCLSTFLQA--AEIFVESELPLSWADRLSGCLRG-----PWAMLVVILAM-LV 691
Db 652 FAICYSALLTNTNCIARIF-----DGVKNGAQRPKFTSPSSQVFCIGLILY 698
QY 692 EVALCTWLVAPPREVVTDMHMLPTE--ALVHCRTSMVSGLAHATNATLAFICFLGT 748
Db 699 QIVWVSVLLILETPG--TRRYTLPEKRETVILKCNVKD--SSMLISLTYDVVLVILCTVYA 755
QY 749 FLVRSQPCYVRARGLTFAMLAFTWVSFPLANOVVLRAVQMGALLC-----801
Db 756 FKTRKCPENENKAKFTGFTMTTCTIIMLAPLPIF---YVSSDYRVQVTTWCISVSLSG 811
QY 802 --VLGIILAAFHLPRCYILMROPGLN 824
Db 812 FVVLGCLFA---PKVHIYVLFQPOKN 833

Search completed: May 19, 2003, 09:50:32
Job time : 16.0911 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:48:15 ; Search time 38.8189 Seconds
(without alignments)
4522.333 Million cell updates/sec

Title: US-09-927-315-15

Perfect score: 4524

Sequence: 1 MCGPAVLGLSTMLLHPGTG.....GPGDAQGQNDGNTGNGKHE 852

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_todent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3255	71.9	858	11	Q923K1	Q923K1 rattus norv
2	3229	71.4	858	11	Q91VA4	Q91VA4 mus musculu
3	3227.5	71.3	858	11	Q925A4	Q925A4 mus musculu
4	3226.5	71.3	858	11	Q925D9	Q925D9 mus musculu
5	3221.5	71.2	858	11	Q925D8	Q925D8 mus musculu
6	3221.5	71.2	858	11	Q923K0	Q923K0 mus musculu
7	1208.5	26.7	840	11	Q920R8	Q920R8 rattus norv
8	1199.5	26.5	842	11	Q920R8	Q920R8 rattus norv
9	1195.5	26.4	842	11	Q925I5	Q925I5 mus musculu
10	1188.5	26.3	842	11	Q923J9	Q923J9 mus musculu
11	1180.5	26.1	842	11	Q923J9	Q923J9 mus musculu
12	1077.5	23.8	763	4	Q8TDJ9	Q8TDJ9 homo sapien
13	1060	23.4	940	13	Q90WU6	Q90WU6 sparus aura
14	1049	23.2	940	13	Q73635	Q73635 fugu rubrip
15	1014.5	22.4	839	4	Q8TE23	Q8TE23 homo sapien
16	1009.5	22.3	877	13	Q9PW88	Q9PW88 carassius a

17	1002.5	22.2	843	11	Q920R7	Q920R7 rattus norv
18	1000.5	22.1	843	11	Q925I4	Q925I4 mus musculu
19	999.5	22.1	843	11	Q923A8	Q923A8 mus musculu
20	933.5	20.6	880	13	Q73639	Q73639 fugu rubrip
21	925	20.4	864	13	Q73637	Q73637 fugu rubrip
22	876.5	19.4	912	11	Q70410	Q70410 mus musculu
23	845	18.7	844	13	Q93552	Q93552 carassius a
24	835.5	18.5	848	13	Q93553	Q93553 carassius a
25	828	18.3	868	13	Q73636	Q73636 fugu rubrip
26	827.5	18.3	856	13	Q73638	Q73638 fugu rubrip
27	760	16.8	875	13	Q73640	Q73640 fugu rubrip
28	616.5	13.6	855	11	Q70409	Q70409 mus musculu
29	612.5	13.5	1218	13	Q90ZP3	Q90ZP3 oncorhynch
30	605.5	13.4	983	11	Q62916	Q62916 rattus norv
31	601.5	13.3	877	4	Q8TBH9	Q8TBH9 homo sapien
32	601.5	13.3	879	11	Q9QY52	Q9QY52 mus musculu
33	592.5	13.1	779	11	Q35269	Q35269 rattus norv
34	583.5	12.9	1156	13	Q98UC6	Q98UC6 gallus gall
35	583.5	12.9	1188	13	Q98UC5	Q98UC5 gallus gall
36	583.5	12.9	1242	13	Q98UC4	Q98UC4 gallus gall
37	575.5	12.7	1199	11	Q9EPV6	Q9EPV6 mus musculu
38	555.5	12.3	866	11	Q35268	Q35268 rattus norv
39	550.5	12.2	977	13	Q9PWE1	Q9PWE1 ictalurus p
40	516.5	11.4	852	11	Q35192	Q35192 mus musculu
41	512	11.3	850	11	Q35189	Q35189 mus musculu
42	507.5	11.2	457	4	Q8TDJ8	Q8TDJ8 homo sapien
43	503	11.1	667	11	Q35267	Q35267 rattus norv
44	503	11.1	768	11	Q35266	Q35266 rattus norv
45	498.5	11.0	803	11	Q35191	Q35191 mus musculu

ALIGNMENTS

RESULT 1

Q923K1	PRELIMINARY;	PRT;	858 AA.
AC	Q923K1:		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Sweet taste receptor TIR3.		
GN	TASIR3.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	STRAIN-WISTAR:		
RA	Neison G., Hoon M.A., Chandrasekar J., Zhang Y., Ryba N.J.P.,		
RA	Zuker C.S.,		
RT	"Mammalian Sweet Taste Receptors."		
RL	Cell 0:0-0(2001).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-WISTAR: TISSUE=CIRCVALLATE PAPILLA;		
RC	MEDLINE=21927605; Pubmed=11917125;		
RA	Li X., Staszewski L., Xu H., Durick K., Zoller M., Adler E.,		
RA	"Human Receptors for Sweet and Umami Taste."		
RT	Proc. Natl. Acad. Sci. U.S.A. 99:4692-4696(2002).		
DR	EMBL; AY032620; AAK51601.1; -		
DR	EMBL; AF456324; AAK10636.1; -		
DR	InterPro; IPR001828; ANF_receptor.		
DR	InterPro; IPR000337; GPCR_Mgr.		
DR	Pfam; PF00003; 7tm_3; 1.		
DR	Pfam; PF01094; ANF_receptor; 1.		
DR	PROSITE; PS00980; G_PROTEIN_RECIP_F3_2; UNKNOWN_1.		
DR	PROSITE; PS50259; G_PROTEIN_RECIP_F3_4; 1.		
KW	Receptor.		
SO	SEQUENCE. 858 AA; 94845 MW; 98890DAF75973B80 CRC64;		
Query Match	71.9%;	Score 3255;	DB 11; Length 858;

Accession	Species	Gene	Protein	Length	MD5	NCBI TaxID	EMBL	GenBank	DBU
D0	305	TSDMLTPNTARQVGLVGFQKALLPERSHYETHALADPAFCASL	-NAELDIEE	363		10090			
QY	365	VVGORPCDCITLTONASAGL	-NHHTFSYAAVVSVAQALNHTLQCNASGCP	416		10090			
D0	364	VMGQRCPRCDDIMQNLSSGLQLVLSAGQHLHQFATYAAVVSVAQALNHTLQCNASHCH		423		10090			
QY	417	AODPYKPMQLLENNYNTLFHFVGLPLRPDSSGVNDMEYDLKLVWQGSVPRLDHVGRENG	476		10090				
D0	424	VEHVLPMQLLENNYNNMFHARDTLTLPDAEGVNDMEYDLKMWQSPPLVLTHTVGFNG	483		10090				
QY	477	SLRTRLTIRHNTSDONKPVARGSCROGEOVRVKGFSHCYDCVDCAGSYRONDDI	536		10090				
D0	484	TLQLQGSMTY--FQNOVPVSGCRQCKDQGVNRVKGFSHCYDCVDCAGSTRKRPDDE	541		10090				
QY	537	ACTFGODEMSPERSTRCFRRSRPFLAMGEPVALLLLLSLALGLVLAALGLFVHRDS	596		10090				
D0	542	TCTPCNOQOWSPKSTACLPFRPFLAMGEPVALLLLLSLALGLVLAALGLFVHRDS	601		10090				
QY	597	PLVQASGPRFLACFGVLCGLVLSVLLFPQPSAPARLQAQPLSHLPPLTGCISTPLQAA	656		10090				
D0	602	PLVQASGGSQFCFFGLICGLFCLSVLLFPGPSSASCLQAQPMANLPPLTGCISTPLQAA	661		10090				
QY	657	EIFESEELPLSMAORLSGCLRGPMANLVLLAMLEVALCTWYLVAEPPEVTYDMHMLPT	716		10090				
D0	662	EFPVESELPLSMAWMLCSYLNGLWAMLVLLATFVEALQAMWILNRPPEVTYDMHMLPT	721		10090				
QY	717	EALVHCRTSRVWSGGLAHATNATLATLCTGLTFLVRSQPCYNRARGLFTAMLTATITWY	776		10090				
D0	722	EVLCHCRVSWVSGILVHITNAMLAFLCPLGTFLVQSPGRYNRARGLFTAMLTATITWY	781		10090				
QY	777	SEVPLLAVVQVLLPAYOMGALLICVGLIAAFLPFCYLLMPOGLNTEPEFLG----	831		10090				
D0	782	SEVPLLAVVQVQAPVOMGAILVCAILVTEFTFCYVLLMPLKLNTEPEFLGRNAK	841		10090				
QY	832	-----GGPGDAQGOND	842		10090				
D0	842	AADENSGGGEAAQCHNE	858		10090				
RESULT 3									
ID	Q925A4	PRELIMINARY;	PRT;	858 AA.					
AC	Q925A4;								
DT	01-DEC-2001 (Tremblrel, 19, Created)								
DT	01-DEC-2001 (Tremblrel, 19, Last sequence update)								
DT	01-JUN-2002 (Tremblrel, 21, Last annotation update)								
DE	Putative taste receptor.								
GN	TIR3.								
OS	Mus musculus (Mouse).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Rodentia; Sclurognath; Muridae; Murinae; Mus.								
OX	NCBI_TaxID=10090;								
RP	[1]								
RC	SEQUENCE FROM N.A.								
RA	STRAIN=129/SV;								
RA	Sains E., Korley J.N., Batley J.F., Sullivan S.L.;								
RT	"Identification of a novel member of the T1R family of putative taste								
RT	receptors.";								
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBU databases.								
DR	EMBL; AY063318; AAK01937.1; -								
DR	InterPro; IPR001828; ANF_receptor.								
DR	InterPro; IPR000337; GPCR_Mgr.								
DR	Pfam; PF00003; 7tm_3; 1.								
DR	Pfam; PF01094; ANF_receptor; 1.								
DR	PROSITE; PS00980; G_PROTEIN_RECPT_F3_2; UNKNOWN_1.								
DR	PROSITE; PS50259; G_PROTEIN_RECPT_F3_4; 1.								
KM	RECEPTOR.								

Query Match	71.3%;	Score 3227.5;	DB 11;	Length 858;
Best Local Similarity	72.4%;	Pred. No. 2.6e-250;		
Matches 620;	Conservative 67;	Mismatches 156;	Indels 13;	Gaps 4;

0Y	5	AVLGLSTLALLHPCTGAPRLCLISQOLRRKGOVDYVGLGRLPLGEAEBAQRLSRTSSPVC	64
0Y	5	AIMGSLIAAFLTEIGMGASLCLISQOFKAGDYLIGLGEPLIDSTEBAUTNORAQNPSTLCNR	64
0Y	65	FSSNGLIALMALMKMAVEEINKSDLLPGLRGLYDLPFTCSEPPVAMKPSLMLFAKGSRD	124
Dd	65	FSPGLGLEFLAMAMKMAVEEINNQSALLRGLRGLYDLPFTCSEPPVYMKSSLMLFAKGSQS	124
0Y	125	IAAYCNTYQYOPRYVLAVIGPHSSSELAATYGFESFEFLMPQVSYGASKEILLASARET	184
Dd	125	IAAYCNTYQYOPRYVLAVIGPHSSSELAATYGFESFEFLMPQVSYGASMDRLSDRETP	184
0Y	185	RTVPEDRQYOLAAAEELILEFEGMNVVAALGSDEEGROGLSFEFSLAARGLCIANHE	244
Dd	185	RTVPEDRQYOLAAAYVTLLOFNSMNVVAALGSDDYGGREGISLFSLSAARGLCIANHE	244
0Y	245	LPRAADSLRGLKQVODYLAQOVNOSSVOYVLLFASVAHAALFNTSISRLSPKVVWAS	304
Dd	245	QHDYSGOOLQGLDYDLRLBOVNOSSKQYVVLLFASARAAYSLFESYIHNGLSPRVWAS	304
0Y	305	TSIDLVMGLRPGMAONGTVLGFLORGAGLOHEFPQYKTHALATDPAFCALSAGEROG	364
Dd	305	TSIDLVTLPNJARQGTVLGFLORGALPERSHYEITHALAADPAFCASLNAEILDEE	363
0Y	365	VVGORPCODCITLLONVASGL-----NNHQTFSYAAYSVAAALNHTLOCNAS	416
Dd	364	VMGORPCODDILMLONSSGLLONLASQOLHQLFATYAAYSVAAALNHTLOCNVSH	423
0Y	417	AODPRKPMQOLLENMTNLTFHYGGLPLRPDSSGANDMEYDLKLMWQSSVRLNHYG	476
Dd	424	VSEHVLPMQOLLENNYMMSEFHARDTLLOFDAGANDMEYDLKMMWQSPRYLHVTG	483
0Y	477	SLRTERLKRIMHTSDNORPVRSCRSOCQEGOVRRYKGFHSQCYDCVCEAGSYKON	536
Dd	484	TLQLOQSKMY--PQNOVPVQSGCNCIDQGVRRYKGFHSQCYDCVCKKSGYKRP	541
0Y	537	ACTFCGODEMSPENSTRCFRRRSRPLMAGEPAYVLLLLLSLALGLYLAALGLFV	596
Dd	542	TCTPCNONOWMSEKSTACLPRRPFLMAGEPVAVLLLLLSLCTVGLALALGLSV	601
0Y	597	PLVQASGGRPLACFGVYCLVCLSVLLEPQGPSPARCLAOPLSHLPLTGCSTLFL	656
Dd	602	PLVQASGGSQCFGLCTGLCTCLSVLPRGRRSSASCLAOQPMANHLPLTGLSTLFL	661
0Y	657	ELFVESELPLSMADRLSGCLRGPMANMLVLLMLVEVALCTWYLVAPREPVYDWM	716
Dd	662	ELFVESELPLSMANMLCYLRLGMLAMLVLLATFEVALCAMLVTLAPPEPVYDWS	721
0Y	717	EALVYCRTRSMVSRGLAATNATLALFCGTFTVRSOPGYNARGLTAMALYFTW	776
Dd	722	EYLEHCHVRSMVSLGLVITNMALEFCFTFTVLOSPGKYNARGLTAMALYFTW	781
0Y	777	SFVPLLANVQVVLRAVNOGALLCYLGLILAAFLHPRCLYLMRPGNLTEBEFLG	836
Dd	782	SFVPLLANVQVQAVNOGAILVCAIGLITVPHLPRCYVLLMLPKINTGGEFLGR	841
0Y	837	AGQGDNGTNGOKHE 852	
Dd	842	AADENS--GGEAAOE 855	

RESULT 4	
ID 0925D9	PRELIMINARY; PRT; 858 AA.
AC 0925D9;	
DT 01-DEC-2001 (Tremblre, 19, Created)	
DT 01-DEC-2001 (Tremblre, 19, Last sequence update)	
DT 01-JUN-2002 (Tremblre, 21, Last annotation update)	
DE Putative sweet taste receptor family 1 member 3.	
CN TAS1R3.	
OS Mus musculus (mouse).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RX MEDLINE=21225287; PubMed=11326277;
 RC STRAIN=C57BL/6J;
 RA Max M., Shanker Y.G., Huang L., Rong M., Campagne F., Liu Z.,
 Weinstein H., Damak S., Margolske R.F.;
 RT "Tas1r3, encoding a new candidate taste receptor, is allelic to the
 RT sweet responsiveness locus Sac.";
 RL Nat. Genet. 28:58-63(2001).
 DR EMBL: AF368024; AAK55536.1; -
 DR MGD; MG1:1933547; Tas1r3.
 DR InterPro; IPR001828; ANF_receptor.
 DR Pfam; PF00003; 7tm_3; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECPT_F3_2; UNKNOWN_1.
 DR PROSITE; PS50259; G_PROTEIN_RECPT_F3_4; 1.
 KW Receptor.
 SQ SEQUENCE 858 AA; 94633 MW; 4E416824FDB478F8 CRC64;

Query Match 71.3%; Score 3226.5; DB 11; Length 858;

Best Local Similarity 72.4%; Pred. No. 3.2e-250;
 Matches 620; Conservative 68; Mismatches 155; Indels 13; Gaps 4;

QY 5 AVLGSLMALHPGTGAPLCLSQOLRMKGVDYVGLGFLPGAEAGLRSTRPSPVCTR 64
 DB 5 AIMGLSLAAFLDGMGASLCLSQOKAGDYIILGFLPGTEATLNGRQPSIPCN 64
 QY 65 FSSNGLLMALAKMAVEINNKSDLLPGLRGLYDLFTCSPPVYAMKPSLMFLAKASRD 124
 DB 65 FSPGLFLAMAKMAVEINNKSGALLPGLRGLYDLFTCSPPVYAMKPSLMFLAKVGSQS 124
 QY 125 IAAACNTQYQPRVLAIVGPHSSSLAWTGKFFSFPLMPQVSYASAMDRLSDRFPSEF 184
 DB 125 IAAACNTQYQPRVLAIVGPHSSSLAWTGKFFSFPLMPQVSYASAMDRLSDRFPSEF 184
 QY 185 RTVPSDRVQTLAAELLQEFGMNVAALGSDDEXRGGLSTFSALAAAGICIAHEGLVP 244
 DB 185 RTVPSDRVQTLAAELLQEFGMNVAALGSDDEXRGGLSTFSALAAAGICIAHEGLVP 244
 QY 245 LPRADSRGLKQVDYLVHQVNSQVYVLLFASVAHAHALFNYSSISRLSPKVVYASAWL 304
 DB 245 QHDSGQGLGVLDYLVHQVNSQVYVLLFASVAHAHALFNYSSISRLSPKVVYASAWL 304
 QY 305 TSDLVMLGPGMAQNGYVGLQKGAOLHEPQYKTHLATATDPAFCSALGEROGLEED 364
 DB 305 TSDLVMLGPGMAQNGYVGLQKGAOLHEPQYKTHLATATDPAFCSALGEROGLEED 364
 QY 365 VVGRCRQCDCITLQNVASGL-----NHQTFSVYAAVYSVAQALHNTLQCNASGCP 416
 DB 365 VVGRCRQCDCITLQNVASGL-----NHQTFSVYAAVYSVAQALHNTLQCNASGCP 416
 QY 417 AADPVKRWOLLENNYNTLFRVGLPLRPDSSGNVDMEDYDKLWYQGSVPRLHVGRENG 476
 DB 417 AADPVKRWOLLENNYNTLFRVGLPLRPDSSGNVDMEDYDKLWYQGSVPRLHVGRENG 476
 QY 424 VSEVLEWQLENNYNNFSFARDLTLOFDAENVDMEDYDKLWYQGSVPRLHVGRENG 483
 DB 424 VSEVLEWQLENNYNNFSFARDLTLOFDAENVDMEDYDKLWYQGSVPRLHVGRENG 483
 QY 477 SLTERLKIMHTSDNCRPVSRCSQOCQOVRRKGFHSCCYDQCEASQONPDI 536
 DB 477 SLTERLKIMHTSDNCRPVSRCSQOCQOVRRKGFHSCCYDQCEASQONPDI 536
 QY 484 TLQLOQSKMTW--PGNVPVYSGCRCKQGVRRVKGFSCTCYCCKASKYKHHDF 541
 DB 484 TLQLOQSKMTW--PGNVPVYSGCRCKQGVRRVKGFSCTCYCCKASKYKHHDF 541
 QY 537 ACTFCGODEMSPERSTRCFRRSRFLAMGEPVALLLLSLALGLVLAALGLFVHHDS 596
 DB 537 ACTFCGODEMSPERSTRCFRRSRFLAMGEPVALLLLSLALGLVLAALGLFVHHDS 596
 QY 542 TCFCPNODQMSPEKSTACLRPRRFLAMGEPVALLLLSLALGLVLAALGLFVHHDS 601
 DB 542 TCFCPNODQMSPEKSTACLRPRRFLAMGEPVALLLLSLALGLVLAALGLFVHHDS 601
 QY 597 PLVQASGGLACGLVGLVGLSVLLFPQGPAPARCLAQOPLSHPLTGLSTLFLQAA 656
 DB 597 PLVQASGGLACGLVGLVGLSVLLFPQGPAPARCLAQOPLSHPLTGLSTLFLQAA 656
 QY 602 PLVQASGGLACGLVGLVGLSVLLFPQGPAPARCLAQOPLSHPLTGLSTLFLQAA 661
 DB 602 PLVQASGGLACGLVGLVGLSVLLFPQGPAPARCLAQOPLSHPLTGLSTLFLQAA 661
 QY 657 EIFVESLPLSMADRLSGCLRGPMVAVLLAMVEVALCWTYVAPPEVYDWMHLP 716
 DB 657 EIFVESLPLSMADRLSGCLRGPMVAVLLAMVEVALCWTYVAPPEVYDWMHLP 716
 QY 662 ETVESESLPLSMANMLCSTLNGMLAVLLATLFAEALCAWYLIAPPEVYDWMVLP 721
 DB 662 ETVESESLPLSMANMLCSTLNGMLAVLLATLFAEALCAWYLIAPPEVYDWMVLP 721

QY 717 EALVHCRTSRWVSFGLAHATNATLAFCLGTFIVRSQPCGNRARGITFAMLAFTW 776
 DB 722 EVLEHCHRSVNSLGVIAHTAMALAFCLGTFIVRSQPCGNRARGITFAMLAFTW 781
 QY 777 SFVPLLANVOYVLRPAVOMGALLICVIGITLAAFLPRCYLLMRQPLNTPPEYAGGEGD 836
 DB 782 SFVPLLANVOYVLRPAVOMGALLICVIGITLAAFLPRCYLLMRQPLNTPPEYAGGEGD 841
 QY 837 AQONGDGTGNGKHE 852
 DB 842 AADENSG--GGEAAOE 855

RESULT 5

QY 0925D8 PRELIMINARY; PRT; 858 AA.
 ID AC 0925D8;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative sweet taste receptor family 1 member 3.
 GN TAS1R3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RX MEDLINE=21225287; PubMed=11326277;
 RC STRAIN=129/SVEV;
 RA Max M., Shanker Y.G., Huang L., Rong M., Campagne F., Liu Z.,
 Weinstein H., Damak S., Margolske R.F.;
 RT "Tas1r3, encoding a new candidate taste receptor, is allelic to the
 RT sweet responsiveness locus Sac.";
 RL Nat. Genet. 28:58-63(2001).
 DR EMBL: AF368025; AAK55537.1; -
 DR MGD; MG1:1933547; Tas1r3.
 DR InterPro; IPR001828; ANF_receptor.
 DR Pfam; PF00003; 7tm_3; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECPT_F3_2; UNKNOWN_1.
 DR PROSITE; PS50259; G_PROTEIN_RECPT_F3_4; 1.
 KW Receptor.
 SQ SEQUENCE 858 AA; 94533 MW; 2012102855D7E6B0 CRC64;

Query Match 71.2%; Score 3221.5; DB 11; Length 858;

Best Local Similarity 72.3%; Pred. No. 8e-250;
 Matches 619; Conservative 67; Mismatches 157; Indels 13; Gaps 4;

QY 5 AVLGSLMALHPGTGAPLCLSQOLRMKGVDYVGLGFLPGAEAGLRSTRPSPVCTR 64
 DB 5 AIMGLSLAAFLDGMGASLCLSQOKAGDYIILGFLPGTEATLNGRQPSIPCN 64
 QY 65 FSSNGLLMALAKMAVEINNKSDLLPGLRGLYDLFTCSPPVYAMKPSLMFLAKASRD 124
 DB 65 FSPGLFLAMAKMAVEINNKSGALLPGLRGLYDLFTCSPPVYAMKPSLMFLAKVGSQS 124
 QY 125 IAAACNTQYQPRVLAIVGPHSSSLAWTGKFFSFPLMPQVSYASAMDRLSDRFPSEF 184
 DB 125 IAAACNTQYQPRVLAIVGPHSSSLAWTGKFFSFPLMPQVSYASAMDRLSDRFPSEF 184
 QY 185 RTVPSDRVQTLAAELLQEFGMNVAALGSDDEXRGGLSTFSALAAAGICIAHEGLVP 244
 DB 185 RTVPSDRVQTLAAELLQEFGMNVAALGSDDEXRGGLSTFSALAAAGICIAHEGLVP 244
 QY 245 LPRADSRGLKQVDYLVHQVNSQVYVLLFASVAHAHALFNYSSISRLSPKVVYASAWL 304
 DB 245 QHDSGQGLGVLDYLVHQVNSQVYVLLFASVAHAHALFNYSSISRLSPKVVYASAWL 304
 QY 305 TSDLVMLGPGMAQNGYVGLQKGAOLHEPQYKTHLATATDPAFCSALGEROGLEED 364
 DB 305 TSDLVMLGPGMAQNGYVGLQKGAOLHEPQYKTHLATATDPAFCSALGEROGLEED 364

```

Db 305 TSDVMTLPNIAVGTVLGFLQRCALPREFSHYETHLADPAFCASL-NAELDLEH 363
QY 365 VVGORCPQDCITLQNVASGL-----NHQTSVYAAYSVAAQALHNTLQCNASCP 416
Db 364 VMGQRCPCDDIMQNLSSGLQNLASQQLHHQJFATYAAYSVAAQALHNTLQCNVSHCH 423
QY 417 AODPKPMQLENNYNTLFHVGGLPLRFDSGNDVMEYDLKLMWQGSVPRLHDVGRNG 476
Db 424 VSEHLPQMLENNYNTMFHARDLTLQFDAGENDVMEYDLKLMWQOSTPVLHVTGTNG 483
QY 477 SLRTERLIRWHTSDNQRPVSRCSQCEGOVRRVKGFSHCYDCVDCAGSYRONPDI 536
Db 484 TLQLOQSKMYW--PQNVQPVSGCSROCKDGOVRRVKGFSHCYDCVDCAGSYRKHDPDF 541
QY 537 ACTPCGODEMSPERSTRCFRRRSRFLANGEPAYLILLLLSLALGLVLAAGLFTVHHDS 596
Db 542 TCTPCNODQMSPEKSTACLPFRPKFLANGEPAYLILLLLSLALGLVLAAGLFTVHHDS 601
QY 597 PLYVAGSGPLACFGLVGLVCLSVLPPGQSPARCIAOQPLSHLPLTGCLSTFLQAA 656
Db 602 PLYVAGSGSQCFCGLICGLICLSVLPPGRSSASCLAOQPMHLPLTGCLSTFLQAA 661
QY 657 EIFVESELPLSMADRLSGCLGPGMAVLVLLAMLEVALCTWYLAPEPEVYTDHMLPT 716
Db 662 ETFVESELPLSMANMLCSYLRGLMAMLVLLATFEVALCAMYLAPEPEVYTDMSVLP 721
QY 717 EALVHCRTSRVSWSGFLATNATLAFCLGFTFLVRSQPCYNRARGLTFMLAFTTW 776
Db 722 EYLEHCHVRSWVSGLVITNMAFLCFLGFTFLVRSQPCYNRARGLTFMLAFTTW 781
QY 777 SEVPLLANVQVAVPAGVOMGALLCVLGIILAFHLPRCYLIMROGLTPEFFLGSGGPD 836
Db 782 SEVPLLANVQVAVPAGVOMGALLCVLGIILAFHLPRCYLIMROGLTPEFFLGSGGPD 841
QY 837 AOGQNDGNTGNOGKHE 852
Db 842 AADENSG--GGEAAQ 855

RESULT 6
Q923K0 PRELIMINARY; PRT; 858 AA.
ID Q923K0
AC Q923K0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Sweet taste receptor TIR3.
GN TASIR3 OR SAC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWR/J;
RA Nelson G., Hoon M.A., Chandrasekar J., Zhang Y., Ryba N.J.P.,
RA Zuker C.S.;
RT "Mammalian Sweet Taste Receptors.";
RL Cell 0:0-0(2001).
DR EMBL; AY032621; AA051602.1; -.
DR MGD; MGI:1933547; Tasir3.
DR InterPro; IPR001828; ANF_receptor.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PROSITE; PS00980; G_PROTEIN_RECEP_F3_2; UNKNOWN.1.
DR PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.
KW Receptor.
SQ SEQUENCE 858 AA; 94530 MW; 686A7E524267796A CRC64;

Query Match 71.2%; Score 3221.5; DB 11; Length 858;
Best Local Similarity 72.4%; Pred. No. 8e-250;
Matches 620; Conservative 67; Mismatches 156; Indels 13; Gaps 4;

```

```

QY 5 AVLGSLMALHPGICAPLCLSQQLRMKGDVYGLPFLGEAEAGLRSRTPSPVCTR 64
Db 5 AIMGSLIAFLELGGASCLISQCFKAQGDYILGLFPLGSTTEATLNRQPNISLQNR 64
QY 65 FSSNGLIWAALMAKVAEINNKSDLLPGLRGLYDLPDCSEPVYMKPSLMEFAAGSRD 124
Db 65 FSPGLGLFLAAMKVAEINNKSSALLPGLRGLYDLPDCSEPVYMKPSLMEFAAGSOS 124
QY 125 IAAYCNTRYQYPRVLAIVGPHSSSELAMTGRFFSEFLMPQVSYGASMEILSARETFPSFF 184
Db 125 IAAYCNTRYQYPRVLAIVGPHSSSELATLTKFFSEFLMPQVSYGASMDRLSDRETFPSFF 184
QY 185 RTVPSDRVQLAAAEELLQDFGNWVAALGSDDEYGRQSLSTFSALAAARGICIAHEGLVP 244
Db 185 RTVPSDRVQLAAAVTLTLQNFMSNMVAALGSDDGREGELSTFSALNARGICIAHEGLVP 244
QY 245 LPRADDSRLGKQVDVLAHVQVSVQVYLLFASVHAHLFNYSSIRLSPKRWVAEAML 304
Db 245 QHDTSGQDLGKVDVLCQVQVSVQVYLLFASARAVYLSFYSIHGSLSPKRWVAEAML 304
QY 305 TSDVMTLPNIAVGTVLGFLQRCALPREFSHYETHLADPAFCASL-NAELDLEH 363
Db 305 TSDVMTLPNIAVGTVLGFLQRCALPREFSHYETHLADPAFCASL-NAELDLEH 363
QY 365 VVGORCPQDCITLQNVASGL-----NHQTSVYAAYSVAAQALHNTLQCNASCP 416
Db 364 VMGQRCPCDDIMQNLSSGLQNLASQQLHHQJFATYAAYSVAAQALHNTLQCNVSHCH 423
QY 417 AODPKPMQLENNYNTLFHVGGLPLRFDSGNDVMEYDLKLMWQGSVPRLHDVGRNG 476
Db 424 VSEHLPQMLENNYNTMFHARDLTLQFDAGENDVMEYDLKLMWQOSTPVLHVTGTNG 483
QY 477 SLRTERLIRWHTSDNQRPVSRCSQCEGOVRRVKGFSHCYDCVDCAGSYRONPDI 536
Db 484 TLQLOQSKMYW--PQNVQPVSGCSROCKDGOVRRVKGFSHCYDCVDCAGSYRKHDPDF 541
QY 537 ACTPCGODEMSPERSTRCFRRRSRFLANGEPAYLILLLLSLALGLVLAAGLFTVHHDS 596
Db 542 TCTPCNODQMSPEKSTACLPFRPKFLANGEPAYLILLLLSLALGLVLAAGLFTVHHDS 601
QY 597 PLYVAGSGPLACFGLVGLVCLSVLPPGQSPARCIAOQPLSHLPLTGCLSTFLQAA 656
Db 602 PLYVAGSGSQCFCGLICGLICLSVLPPGRSSASCLAOQPMHLPLTGCLSTFLQAA 661
QY 657 EIFVESELPLSMADRLSGCLGPGMAVLVLLAMLEVALCTWYLAPEPEVYTDHMLPT 716
Db 662 ETFVESELPLSMANMLCSYLRGLMAMLVLLATFEVALCAMYLAPEPEVYTDMSVLP 721
QY 717 EALVHCRTSRVSWSGFLATNATLAFCLGFTFLVRSQPCYNRARGLTFMLAFTTW 776
Db 722 EYLEHCHVRSWVSGLVITNMAFLCFLGFTFLVRSQPCYNRARGLTFMLAFTTW 781
QY 777 SEVPLLANVQVAVPAGVOMGALLCVLGIILAFHLPRCYLIMROGLTPEFFLGSGGPD 836
Db 782 SEVPLLANVQVAVPAGVOMGALLCVLGIILAFHLPRCYLIMROGLTPEFFLGSGGPD 841
QY 837 AOGQNDGNTGNOGKHE 852
Db 842 AADENSG--GGEAAQ 855

RESULT 7
Q920R8 PRELIMINARY; PRT; 840 AA.
ID Q920R8
AC Q920R8;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Putative taste receptor Trl (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```

OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MISTAR;
 RX MEDLINE=99159821; PubMed=10052456;
 RA Hoon M.A., Adler E., Lindemeier J., Battey J.F., Ryba N.J.,
 RA Zuker C.S.;
 RT "Putative mammalian taste receptors: a class of taste-specific GPCRs
 RT with distinct topographic selectivity.";
 RL Cell 96:541-551(1999).
 DR EMBL: AF127389; AAD18069.1;
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCR_MGR.
 DR PROSITE: PS50259; G_PROTEIN_RECPT_F3_4; 2.
 KW Receptor.
 FT NON_TER
 SQ SEQUENCE 840 AA; 93496 MW; 1FCBF7EFC6B845DB CRC64;

Query Match 26.7%; Score 1208.5; DB 11; Length 840;

Best Local Similarity 33.8%; Pred. No. 3e-88; Matches 291; Conservative 151; Mismatches 339; Indels 79; Gaps 18;

QY 12 WALHPGTGAPLCLSQL-----RMKGDYVLGGLFPL-GEAEAGLR 53
 DB 4 WA-----AHLLSLQLVYCAFCQRTSSPGFSLPGDFLLAGLSHGD---LQV 52
 QY 54 RTRSSPYCVR---FSSNGLLMALAMKAVEINNKSDLPGLGLDLPFCSEPVYAM 110
 DB 53 RHRPLVTSQDPSDFNGHGLFOAMRFTVEINNSLLPNTLGLYELXVCSES-ANV 111
 QY 111 KPSIMFLAKAGSRDIAYCNYTOYQPRVLAIGPHSELAWYTKFSEFLMPQVSGAS 170
 DB 112 YATIRVLAIGPRHIELOKDLRNHSSKVAATIGPDNDTHATTAALGFLMPPLVYSAS 171
 QY 171 MELISARETPSPFRIVPSDRVQLTAAAEILOEFGMMVVAALGSDDEYRGGLSIFSALA 230
 DB 172 SVVLSAKRKPSFLRTVPSDRHVEVWVQLQSFQGWVWISLIGSGYGOGLVQALEELA 231
 QY 231 AARGICIAHEGLVPLPRADSRIG--KYQDVLIHOVNOSSVQVVLAFASVHAHAFNFSI 288
 DB 232 VPRICVAFKQIVP---SARVSDPRQSMOHLAQTIVVVSFNRHLARVEFSNV 287
 QY 289 SSRSPYVWVAASEAMLTSDLVMGRLPGMAOMQTVLGLFQGA---LHEFP-QYKTHLAL 344
 DB 288 LANLIGKVVWASEDMALISTYITSTVGIGTIVGLVAVQQRQVGLKFEFESYVRAVYAA 347
 QY 345 AT---DPAFCGALGERDEGLEDVVGORPCQDCITLQNS--AGLNHHTFSVYAAVYS 399
 DB 348 PSACPESGWS---TNOICRECHFTTRNMPTLGAFSMAAYRUYEAUYA 394
 QY 400 VAAQLNHTLQCNASGCAOPDKVPMOLLENNYLNLFHVGLRLPFRSSGNDMEYDLKMW 459
 DB 395 VAGHLHOLLGCTSTIC--SRGVPYPMQLQYKVFNLHENTVAFDDGDLGYDDITLAW 453
 QY 460 VMQGSVPLHDVGRFNS---LATERLKIRWHTSDNOKPVSRCSROCEQGVRRYKGFHS 516
 DB 454 DMNRPWTPELIGSASLPHLDINKKIOHGNKNQVPVSVCTDCLAGHNRVVGSHH 513
 QY 517 CNYCVCEAGSYRQNDIACFTGODEMSPSTRCFRRRSRFLAMGERAVALLLL 576
 DB 514 CCEFCVCEAGTFLNMEELHICQPCGTEWAPKSESTCFPRVTEFLAMHEPISLVLTAAN 573
 QY 577 SLALGVLAALGLFVHRDPSLVQASGGLACFGLVGLVLCVSLLEFPQSPARCLAQ 636
 DB 574 TLLLLLVGTRAGLRAMHFRPVVASAGRCETMLGSLVAGSCGFYFGEPTVPACILR 633
 QY 637 QPLSLPLVGLSLTLFLQAAEIFV---ESELPL---SWADRLSGCLRGPAAMLVLILAM 689
 DB 634 QPLSLSLFAIFLSGLTIRSPQVLIIFFKSTIKVPIFFYRTWA-----QNHGAGLFVYSS 686

QY 690 LVEVALCTWYLVAEPPEVYTDWMLPTLEALVHCRTSRWSVFGLAHATNATLAFCLGTF 749
 DB 687 TVHLILCTLVVMWTPPTREYQRFPHVLILECTEVNSVGFLLAFNHLISLTFVCSY 746
 QY 750 LVRSQPCYRARRBLTAMALAFITWSEFPLANOVLRPAVQKALLLCVIGILAAF 809
 DB 747 LGKELPENYENAKCVTSFLINFEVSWIAFTMASIYGOSYLPANVLAGLTTLSGFSGY 806
 QY 810 HLPKCYILMRQPLNTEPEEF 829
 DB 807 FLPKCYIILCRPELNTNEH 826

RESULT 8

ID 099PG5 PRELIMINARY; PRT; 842 AA.
 AC 099PG5;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
 DE Putative sweet taste receptor TIR1 (Fragment).
 GN TASI1R OR GPR70.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129P3/J;
 RX MEDLINE=21030739; PubMed=11178737;
 RA Li X., Inoue M., Reed D.R., Hugue T., Puchalski R.B., Tordoff M.G.,
 RA Nimolaya Y., Beauchamp G.K., Bachmanov A.A.;
 RT "High-resolution genetic mapping of the saccharin preference locus
 RT (Sac) and the putative sweet taste receptor (TIR1) gene (Gpr70) to
 RT mouse distal chromosome 4.";
 RL Mamm. Genome 12:13-16(2001).
 DR EMBL: AF301162; AAK07092.1;
 DR MGD: MGI:1927505; Tas1r1.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000345; Cytochrome_bind.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCR_MGR.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS50259; G_PROTEIN_RECPT_F3_4; 1.
 KW Receptor.
 FT NON_TER
 SQ SEQUENCE 842 AA; 93428 MW; 39739A2FF482D33F CRC64;

Query Match 26.5%; Score 1199.5; DB 11; Length 842;

Best Local Similarity 33.2%; Pred. No. 1.6e-87; Matches 285; Conservative 159; Mismatches 340; Indels 75; Gaps 17;

QY 12 WALHPGTGAPLCLSQL-----RMKGDYVLGGLFPLGAEAEGLR 52
 DB 4 WA-----AHLLSLQLVYCAFCQRTSSPGFSLPGDFLLAGLSHGD---HADCLQ 53
 QY 53 SRTSPSPYCVCR---FSSNGLLMALAMKAVEINNKSDLPGLGLDLPFCSEPVYAM 110
 DB 54 VHRPLVTSQDPSDFNGHGLFOAMRFTVEINNSLLPNTLGLYELXVCSESNNV 113
 QY 108 VAMPKSLMFLAKAGSRDIAYCNYTOYQPRVLAIGPHSELAWYTKFSEFLMPQVSY 167
 DB 114 YV---TLRVLAQOSTGHELMQDLRNHSSKVAALIGPDNDTHATTAALLSPFLMPVSY 170
 QY 168 GASMLISARETPSPFRIVPSDRVQLTAAAEILOEFGMMVVAALGSDDEYRGGLSIFS 227
 DB 171 EASSVILSGKRKPSFLRTVPSDRHVEVLRVQLQSFQGWVWISLIGSGYGOGLVQALE 230
 QY 228 ALAARGICIAHEGLVPL-PRADSRIGKYQDVLIHOVNOSSVQVVLAFASVHAHAFN 286
 DB 231 ELATPRGICVAFKQIVPLSLAQAGDPRM---QRMALRLARATTVVVSFNRHLARVEFSNV 287

DE Taste receptor T1R1.
 GN TAS1R1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nelson G., Hoon M.A., Chandrasekar J., Zhang Y., Ryba N.J.P.,
 RA Zuker C.S.;
 RT "Mammalian Sweet Taste Receptors";
 RL Cell 0:0-0(2001).
 DR EMBL: AY032622; AKS1603.1; -.
 DR MGD: MGI:1927505; Tas1r1.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS50259; G_PROTEIN_RECPT_F3_4; 1.
 KM Receptor.
 SQ SEQUENCE 842 AA; 93455 MW; D4D2CE90959E991A CRC64;

Query Match 26.3%; Score 1188.5; DB 11; Length 842;
 Best Local Similarity 33.1%; Pred. No. 1.2e-86;
 Matches 284; Conservative 159; Mismatches 341; Indels 75; Gaps 17;

12 WALHPTGAPICLSOOL-----RMKGDYVIGLFFPLGEAEAGLR 52
 4 WA-----HLLSLQLAAVAYCAFSQORTSSPGFSLPEDFLAGLFSL---HADCLQ 53
 53 SRTSPSSPVCTR---FSSNGLLMALAMKAVEEINNKSDLLPGRLGIDLEFDTCSER--V 107
 54 VRHRLPVTSCDRSDSFNGHGHLEFQAMRFYEEINNTALLPNTITGLAYLVCSESSNV 113
 108 VAMKPSLFLAKAGSRDIAATCANTQYQPRVLAIVGHSSELAVYTKFSFELMPQVSY 167
 114 YA---TLRVPAOQGTGHEMORDLRNHSKRVALLIGPNTDHAVTALLSPFLMPVSY 170
 168 GASMEILSARETFPSFRTVPSDRVOLTAALAELOEFGMMNVALGSDDEXGROGLSIFS 227
 171 EASSVILSGKKRPFSLTIPSDKYQVEVYRLLSQSGFWWISLVGSDGIGQLGVQALE 230
 228 ALAARGICIAHEGLVPL-PRADSRIGKYQVLDHOVQSSQVVLLEFASVHAHALENY 286
 231 ELAPRGICVAFKDYVPLSAQADPRM---QRMALRLARATVYVVESSNHLGVFERS 287
 287 SISRSLSPKRVVASEANLTSVLVGLPGMAQMGTVLGLQAGDLHERPQVYKTHL-ALA 345
 288 VVLANLIGKVIASEDMAISTYIYNVPGIOIGTVLGAVALQDQVPGIKFEESYVQAVM 347
 346 TDPAPC---SALGREGGLEEDVVGORCPQDCITLQNV--AGLNHQTSVAAAVSV 400
 348 GAPRTCEBGSNG-----TNQLCRECHATTNNMPLGAFMSAAVNYEAYTAV 397
 401 AQAHLNLTQCNASGCPADDPYKPMQLLENMYNLTFHVGGLEPLRDSGNDVMEYDLKLV 460
 398 AHGHQLIGCTSGTC-ARGPYVPMQLQIYKVFLLHKKTVAFDDGDDPLGYDIAMD 456
 461 WQGVSPRLHDVGRNGS---LTERLAKIRWTSNOKRVSCSQCOCQGVRRYKGFHSC 517
 457 WNGEMTFEYVYASLSLSPVHLIDINKTKIOWGNKNNOVAVSVCTRDCEGHRLVWGSNHC 516
 518 CYDQVDEAGSYRONPDIACTPCGODEMSPERSTCFRRRSRLANEPVILLILLIS 577
 517 CFEBMPCEAGFELNTSELHTQPCGTEEMAEBSGACFSRVEELGWEPISLVLLANT 576
 578 IALGLVLAALGLFVHNHSDPLVQASGGLACFGLVLCVLSVTLFPGQSPARACQAO 637
 577 LLLLLLIGTAGLFWMLRHTFVVRBSAGRLCLMLGSLVAGCSLXSPFGKTYVACILRQ 636
 638 PLSHLPLTGLCLSTLQLAAELFV---ESSELPL--SWADRLSGCLRGPMAMLVVLLAML 690

DB 637 PLFSLGFAIFLSCLTIRSFOLVIFKFKSTKVPTEYHTWA-----ONHGAGIFVIYST 689
 QY 691 VEVALCTWVLYVAFPEPVYVOMHMLPTBALVHCRRSVNSGCLANATATLAFCLCTFL 750
 DB 690 VHEFLCTWLAAMWPTRETRORFPHLVILECTEVNSVGLVAFANHLISISTFVCSYL 749
 QY 751 VRSQPGCYNRARGLTFAMLAVFITWSEFVPLANVQVLRPAVOMGALLCVLGILAAH 810
 DB 750 GKELPENYNEKCVTFSLHLHFVSWIAFFTMSSIQSGSYLPANVAVLAGLATLSCGSGYF 809
 QY 811 LPRCYLLMRQGLNTPPEF 829
 DB 810 LPKCYVILCRPELNTEHF 828

RESULT 11
 099PG6 PRELIMINARY; PRT; 842 AA.
 AC 099PG6;
 DT 01-JUN-2001 (TREMURel. 17, Created)
 DT 01-JUN-2001 (TREMURel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMURel. 21, Last annotation update)
 DE Putative sweet taste receptor T1R1 (Fragment).
 GN TAS1R1 OR GPR70.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6BYJ;
 RX MDLINE=21030739; PubMed=11178737;
 RA Li X., Inoue M., Reed D.R., Hugue T., Puchalski R.B., Tordoff M.G.,
 RA Nimonyi Y., Beauchamp G.K., Bachmanov A.A.;
 RT "High resolution genetic mapping of the saccharin preference locus
 (sac) and the putative sweet taste receptor (T1R1) gene (Gpr70) to
 mouse distal chromosome 4.";
 RL Mamm. Genome 12:13-16(2001).
 DR EMBL: AF301161; AKR07091.1; -.
 DR MGD: MGI:1927505; Tas1r1.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRGR.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS50259; G_PROTEIN_RECPT_F3_4; 1.
 KM Receptor.
 FT NON_TER
 SQ SEQUENCE 842 AA; 93425 MW; 58826CA3F5D352E CRC64;

Query Match 26.1%; Score 1180.5; DB 11; Length 842;
 Best Local Similarity 32.9%; Pred. No. 5.3e-86;
 Matches 283; Conservative 159; Mismatches 342; Indels 75; Gaps 17;

12 WALHPTGAPICLSOOL-----RMKGDYVIGLFFPLGEAEAGLR 52
 4 WA-----HLLSLQLAAVAYCAFSQORTSSPGFSLPEDFLAGLFSL---HADCLQ 53
 53 SRTSPSSPVCTR---FSSNGLLMALAMKAVEEINNKSDLLPGRLGIDLEFDTCSER--V 107
 54 VRHRLPVTSCDRSDSFNGHGHLEFQAMRFYEEINNTALLPNTITGLAYLVCSESSNV 113
 108 VAMKPSLFLAKAGSRDIAATCANTQYQPRVLAIVGHSSELAVYTKFSFELMPQVSY 167
 114 YA---TLRVPAOQGTGHEMORDLRNHSKRVALLIGPNTDHAVTALLSPFLMPVSY 170
 168 GASMEILSARETFPSFRTVPSDRVOLTAALAELOEFGMMNVALGSDDEXGROGLSIFS 227
 171 EASSVILSGKKRPFSLTIPSDKYQVEVYRLLSQSGFWWISLVGSDGIGQLGVQALE 230

RA Flanagan J.A.;
 RT "Calcium sensing receptor:cDNA sequence and mRNA expression in a
 marine teleost."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ289717; CAC41352.1;
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PROSITE: PS00979; G_PROTEIN_RECIP_F3_1; UNKNOWN_1.
 DR PROSITE: PS50259; G_PROTEIN_RECIP_F3_4; 1.
 KM Receptor.
 SQ SEQUENCE 940 AA; 105455 MW; 19C3CEDC75A0E3F9 CRC64;

Query Match 23.4%; Score 1060; DB 13; Length 940;
 Best Local Similarity 30.4%; Pred. No. 2,9e-76;
 Matches 26; Conservative 150; Mismatches 341; Indels 118; Gaps 26;

QY 27 QQLRMKGDYVGLFPLGEAEAGLSR-----TRSSPVCTRFSSNGLLMALMKMAVE 81
 DB 24 QRAQMTGDILGLFPI-----HFGVASKDDDLAARPESSOCVRENFRGFRMLQAMIFAID 79

QY 82 EINKSKDLLPGLRLGYDLFDTCSEPVYAMKPSLMFLA--KAGSRDIAAYCNYQYQRYL 139
 DB 80 EINKSSTLPLRITLGYRIFDTCNTVSKALEATLSFVAQNKIDSLNIDFECTDHIPATY 139

QY 140 AVIGPSSSELAMWTGKFFSFLMPQVSYGASMEELISARETFSPFRTVPSDRVQLTAAE 199
 DB 140 AVVGAAGSSTAVANLLGLFYLPIQISVASSRLLSNKQKSMRMLPTDEYQATMAD 199

QY 200 LLOFGNNWVAALGSDDEYGRQGISALAAAGCTIAHGGVLPRLRADSRIGKYQDV 259
 DB 200 IIEFGNNWVASAVASDDYGRPVGEKREKEMEEDICIHNLISQYFEDH---ETQAL 255

QY 260 LHVGNSSQVVLIFASVHAHALFNYSISRLSPKTVVASEAMVLS-----DLVM 310
 DB 256 ADRIENSTAKYIVFAASGPDIETIKEMVRNNTIDRIWLSAEMSSSLIAKPEYLDVYA 315

QY 311 GLPQMA-QMGTVLGFLRGAQLHEFPQYVKT-----HLAATDPAFCS 352
 DB 316 GTIGFALKAGHIPIGF-----REFLOQVQPKDSHNEFVREFMEFNYCLEDSF--- 364

QY 353 ALGRREGGL-----EEVVGQRPCDCITLQNVASAGLNHQTFSSYAAYVSAQA 403
 DB 365 RLQSENGSSTFRPLCTGEEDITSEVPYLDYTHLR-----ISYNYVAAYVSAQA 415

QY 404 LHNFLQNA-----SGCPADQVPRKPMOLLENNYNTLF--HVGGLPLRPDSSGVAMEYD 455
 DB 416 LQDILCTPGOGLEFANNSCADIKKEMAKQVYKQLRHLNTYNSMGEKMFEDNADLAANTY 475

QY 456 LKLM---WMOGSVPRLDHVGFRN-----GSLRTERLKIHWTSDNQKPVSRCSROCEG 506
 DB 476 LNNHRSAGDSQSV--FEVGYNNHAKRGKLFIDNTKILMNGFSSEVPSCNCEDEBP 534

QY 507 QVRR-VAGFSHCCTCYDCVCEAGSTRQNPDIACITFCGODEMSPRSRTRCFRRSRFLAMG 565
 DB 535 TRKGIIDSMPTCECECECSGDESYHKDASVCTKCPNNSNNGNHTFCFELKEITFLSWT 594

QY 566 EPAVILLILLISLALGLVIA--LGLFVHHRSDPLVOASGGPLA---CFGVYGLVCL- 619
 DB 595 EP--FGIALAICAVLAVGLVLAFAVGVFRFRNPIYKATINRELSYVLLFSLIC---CFS 648

QY 620 SVLLFPGQSPARCLAOQPLSHLPLTGLSTLFLQAAEIFV--ESELPLSWADRLSGCLR 677
 DB 649 SSLIFISQPDQMMCRMLQPAFGISFVLCISGLYKTRVLLVFEAKIPTS-----LH 700

QY 678 GPM-----ANLVILLAMLEVAALCTWTLVAPPEVNTDMHMLPTREALVHCRTSRVSGFL 732
 DB 701 RKMWGLNLOGLFVAVLCTFVQVMICVWVLYNAPSSYNN--HDIEIIFITCNEGSVMAIGF 759

QY 733 AHATNATLAFCLGTFGLVVSQPCYNRARGLTFAMLAFTTWVSFPLANQVVLARA 792
 DB 760 LIGYTCLLAICFEPFAKSRKLPENFTFAKFTFSMLIFLWISFLPAVFTYKGFVSA 819

QY 793 VOMGALLCVGLIAAFHLRPLCYLLMRQPLNTP 827
 DB 820 VEATAILASSFGMLACIFKFNKVIILLRPSRNTIE 854

RESULT 14
 ID 073635 PRELIMINARY; PRT; 940 AA.
 AC 073635;
 DT 01-AUG-1998 (TEMBLrel. 07, Created)
 DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)
 DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
 DE Calcium2+ sensing receptor.
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Takifugu.
 OX NCBI_TaxID=31033;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9826788; PubMed=9560249;
 RA Naito T., Saito Y., Yamamoto J., Nozaki Y., Tomura K., Hazama M.,
 RA Nakashishi S., Brenner S.;
 RT "Putative pheromone receptors related to the Ca2+-sensing receptor in
 fugu."
 RT Proc. Natl. Acad. Sci. U.S.A. 95:5178-5181(1998).
 RL EMBL: AB008857; BAA26122.1; -
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR InterPro: IPR000651; RasGEF.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMRG.
 DR PROSITE: PS00979; G_PROTEIN_RECIP_F3_1; 1.
 DR PROSITE: PS50259; G_PROTEIN_RECIP_F3_4; 1.
 KM Receptor.
 SQ SEQUENCE 940 AA; 105814 MW; 06DAB7803B6878B3 CRC64;

Query Match 23.2%; Score 1049; DB 13; Length 940;
 Best Local Similarity 29.5%; Pred. NO. 2.2e-75;
 Matches 255; Conservative 157; Mismatches 357; Indels 96; Gaps 21;

QY 26 SOQLRMKGDYVGLFPLGEAEAGLSR-----TRSSPVCTRFSSNGLLMALMKMAV 80
 DB 24 NQRAQMTGDILGLFPI-----HFGISSKDENLAAARESTKYAFNRGRFRMLQAMVFAI 79

QY 81 EINKSKDLLPGLRLGYDLFDTCSEPVYAMKPSLMFLA--KAGSRDIAAYCNYQYQRYL 138
 DB 80 EINKSSTLPLRITLGYRIFDTCNTVSKALEATLSFVAQNKIDSLNIDFECTDHIPAT 139

QY 139 LAVIGPSSSELAMWTGKFFSFLMPQVSYGASMEELISARETFSPFRTVPSDRVQLTAA 198
 DB 140 IAVVGAAGSSTAVANLLGLFYLPIQISVASSRLLSNKQKSMRMLPTDEYQATMAD 199

QY 199 ELLOFGNNWVAALGSDDEYGRQGISALAAAGCTIAHGGVLPRLRADSRIGKYQDV 258
 DB 200 DVLEFPQNNWVAVASDDYGRPGIETKFEKEMEERDICHNLISQYFED---CEIKA 255

QY 259 VLVHGNSSQVVLIFASVHAHALFNYSISRLSPKTVVASEAMVLSDLVGLPGMAQ 318
 DB 256 LVRIENSTAKYIVFAASGPDIETIKEMVRNNTIDRIWLSAEMSSSLIAKPEYLDV 315

QY 319 GTVIGFLRGAQL--HEFPQYVKT-----HLAATDPAFCSALGERE 358
 DB 316 EGTIGFVLKAGNIPGFRREFLOQVQPKRGSNEVREFMEFTNYCLEDSF---RLQSE 371

QY 359 QG-----LEEDVVGQRPCDCITLQNVASAGLNHQTFSSYAAYVSAQAQLHNTLQ 409
 DB 372 NGSISFRPLCTSEEDITSEVPYLDYTHLR-----ISYNYVAAYVSAQAQLDILS 422

QY 410 CMA-----SGCPADQVPRKPMOLLENNYNTLF--HVGGLPLRPDSSGVAMEYDLKLM-- 459

```

Db 423 CTPHGGLFANNSCADIKKEMAOYLKQRLHNTYNSMGEKXHFEDNADMEANTYITINHR 482
Qy 460 -WMOGSPRLHDVG-----RPNGLRTERLKIRNHTSDNOKPVSRGCOEGOVRR-V 511
Db 483 SAEGSV-AFREVGIVHNAHARGAKLLIDNTRKMMNNAVSSEVPSCSEDEPGTRKGI 541
Qy 512 KGFHSCCDYDCEAGSRONDDIACTEGODEMSPERSRRCFRRRRFLAMGEPAYLL 571
Db 542 DSMPTCECECTECSDGEISDKHDSICTKCPNNKSSGNHTFCFLKETEFLAMSDEPGIA 601
Qy 572 LLLSLALGLVLAALGLFVHHRDSPLYOASGGLACGLVCLVCL-5VLLFPGPS 629
Db 602 LAICAVIGLVLLAFVAGVFRNMPYKASNRELST-VLLSLICFSSLLIFIGEPQ 659
Qy 630 PARCIAOQPLSHLPLTGCLSTFLQAAETIV--ESELPLS-----WADRLSGCLRGPMW 682
Db 660 DMTGRLRQPAIGISFVLCISCIIVKTRVLLVEFAKIPISIRKRWGINTLQ-----F 711
Qy 683 LVVLLAMLEVALCTWYLVAFRPEVVDHMLPFEALVHCSTRSVSGFLAHATNATLAF 742
Db 712 LVLVLCFTEVQVVICVWLYNAPSSYRN-HDIDETITTCDEGSMVALGFLIGYCLIAA 770
Qy 743 LCLFGLTVRSQPCGYNRARGLTFAMLAFTTWVSFVPLANVOYVLRPAVQMGALLICV 802
Db 771 ICFFPAFRSKLPENFTFAKFTFCMLIFFTWISFIPAYSTYKFSVAEALAILASS 830
Qy 803 LGITAAFLPRCYLLMRQPLNTPR 827
Db 831 YGMALCIFENKRYIILFKRCRNTIE 855

```

RESULT 15

```

ID 08TE23 PRELIMINARY; PRT; 839 AA.
AC 08TE23;
DT 01-JUN-2002 (TREMUREL 21, Created)
DT 01-JUN-2002 (TREMUREL 21, Last sequence update)
DT 01-JUN-2002 (TREMUREL 21, Last annotation update)
DE Taste-specific G-protein coupled receptor T1R2.
CN T1R2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21927605; PubMed=11917125;
RA Li X., Staszewski L., Xu H., Durick K., Zoller M., Adler E.;
RT "Human receptors for sweet and umami taste."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4692-4696(2002).
DR EMBL: AF458154; AAM12239.1; JOINED.
DR EMBL: AF458150; AAM12239.1; JOINED.
DR EMBL: AF458151; AAM12239.1; JOINED.
DR EMBL: AF458152; AAM12239.1; JOINED.
DR EMBL: AF458153; AAM12239.1; JOINED.
KW Receptor.
SO SEQUENCE 839 AA; 95072 MW; 8CEBD6F1503F05FDE CRC64;

```

Query Match 22.4%; Score 1014.5; DB 4; Length 839;
 Best Local Similarity 30.7%; Pred. No. 1.1e-72;
 Matches 264; Conservative 152; Mismatches 378; Indels 67; Gaps 22;

```

Qy 2 IGPAY-----IGLSMALIHETGAPLCISQQLRMKGDYVLCGLFPLGAEAGLRSKR 56
Db 1 MGPRAKTTCSLEFLLMAEPANSDFYL-----PGDYLLGGLFSL-HANMKGIYHLNF 53
Qy 57 PSSPYCTFRSSNGILMAL--AKMAVEEINNKSDLLPGLRGYDLFDC--SEPVVAMKP 112
Db 54 IQVPCKEYEVKIVGINLMQAMRFAVEEINNDSSLLPGVLLGYEIVDYCISSNV---QP 110
Qy 113 SLMLAKGSRDIAYCNYTOYQPRVLAIVGHSSELAMVYTGKFRSFLMPQVSYGASME 172

```

```

Db 111 VLYFLAHNDNL-LPQDEYSNYSIRVAVIGDENSESVMTVANFLSLFLLPQITYSALSD 169
Qy 173 LISARETPPSFFRYVSPDRVOLTAALAEILOEGNMVVALSDDEYGGQGISFISALAA 232
Db 170 ELRQKVRREPALLRTPPSADHVEANVOQLMHERMMIIVLVSSDYTGDNQGLGERVAR 229
Qy 233 RGICIAHGLVPLPRADDSRLG---KVQDVLHQVNOSSVOVLLFAVHAHALEFNYSI 288
Db 230 RDICIAFETLPTLPQNMNMSEERQRLVITYDKLQOSTARVYVFSDDLTIYHFENEVL 289
Qy 289 SSRSPKVVAVSEAMITDYLWGLRPGMAOMGTVLGFLQAGALHFE-----PQYKTH 341
Db 290 RQNFQAVWIAISESAIDPVLHNLDELHGLGFLGITIQSVIIPFSEPEREMGPAGPFP 349
Qy 342 LALATDPFCSALGEREGLEDVYGRCPQDCITTLQ-NVSAGLNHQ-TFSYVAAYS 399
Db 350 LSRTSOSYTCN-----QECDCNLANTLSFNTLLRISGERVYVSVSAYYA 394
Qy 400 VAQALHNTLOCNAGCPADDPYKPMQLENNYNTLFHYGGFLPRDSSGNVMEYDLKLM 459
Db 395 VAHALSHSLGCDKSTC-TRKVVYPMQLEBKVNFLLDHOIFDPQGDVALHLEYQW 453
Qy 460 WMOGSPRLHDVGRENGSLRTERLK----IRNHTSDNOKPVSRGCOEGOVRRYKGFH 515
Db 454 QMDRSQNFQSVASYYPQOR--QLKNIDISHTVNTNTPMSMCKRCSGQKKRPVGIH 511
Qy 516 SCCTCYDCEAGSY-RONDDIACTEGODEMSPERSRRCFRRRRFLAMGEPAYLLLL 574
Db 512 VCCFECIDCLPCTFLNHTDEDEYCOACPNNEMSYSESCFRLQVLEMHAPPIYVAL 571
Qy 575 LLSLALGLVLAALGLFVHHRDSPLYOASGGLACGLVCLVCLSVL-LFPGPSPARC 633
Db 572 LAALEFLSLIALVLFMHRFQPIYRSAGPR-CLMLLTLVLAAMVYVYVGPVKVSC 630
Qy 634 LAQOPLSHLPLTGCLSTFLQAAETIVESELPVSNADRLSGCLR--GPM--AMLVILIA 688
Db 631 LCRQALFPLCEFTICISCIANVRSFOIVCAFKMASRFPAYSYWRYOGPVSAFVTLVKM 690
Qy 689 MIVEVALCTWYLVAFRPEVVDHMLPFEALVHCSTRSVSGFLAHATNATLAFICPLGT 748
Db 691 VIVVIGMLA---TGLSPTRTDPD-PKITTYSQNPNTNSLNTSLDLSVVGFSFA 746
Qy 749 FLVRSQPCGYNRARGLTFAMLAFTTWVSFVPLANVOYVLRPAVQMGALLICVILIA 808
Db 747 YMGKELPTNYNAKFTILSMTEFTSSVSLCFMNASVGLVITYIDLTVNLALISIG 806
Qy 809 FHLPRCYLLMRQPLNTPRPF 829
Db 807 YGPRCYMLLFEPRNTPAYF 827

```

Search completed: May 19, 2003, 09:52:40
 Job time : 41.8189 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:48:40 ; Search time 14.6201 Seconds
(without alignments)
1714.646 Million cell updates/sec

Title: US-09-927-315-15

Perfect score: 4324

Sequence: 1 MCGPAAVGLSLMALHPGTG.....GPGDAQGQNDGNTGNQKHE 852

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1089.5	24.1	1059	4	US-09-134-513-2
2	1068.5	23.6	1078	1	US-08-485-588-7
3	1068.5	23.6	1078	1	US-08-484-565-7
4	1068.5	23.6	1078	2	US-08-480-751-7
5	1068.5	23.6	1078	2	US-08-943-986-7
6	1068.5	23.6	1078	3	US-08-353-784-7
7	1068.5	23.6	1078	3	US-08-484-719B-7
8	1068.5	23.6	1078	4	US-08-484-159-7
9	1060	23.4	1085	1	US-08-485-588-5
10	1060	23.4	1085	1	US-08-484-565-5
11	1060	23.4	1085	2	US-08-480-751-5
12	1060	23.4	1085	2	US-08-943-986-5
13	1060	23.4	1085	3	US-08-353-784-5
14	1060	23.4	1085	3	US-08-484-719B-5
15	1060	23.4	1085	4	US-08-484-159-5
16	1057.5	23.4	1088	1	US-08-485-588-6
17	1057.5	23.4	1088	1	US-08-484-565-6
18	1057.5	23.4	1088	2	US-08-480-751-6
19	1057.5	23.4	1088	2	US-08-943-986-6
20	1057.5	23.4	1088	3	US-08-353-784-6
21	1057.5	23.4	1088	3	US-08-484-719B-6
22	1057.5	23.4	1088	4	US-08-484-159-6
23	1056	23.3	1027	4	US-09-162-021B-2
24	1054	23.3	1079	1	US-08-485-588-8
25	1054	23.3	1079	1	US-08-484-565-8
26	1054	23.3	1079	2	US-08-480-751-8
27	1054	23.3	1079	2	US-08-943-986-8

28	1054	23.3	1079	3	US-08-353-784-8	Sequence 8, Appl1
29	1054	23.3	1079	3	US-08-484-719B-8	Sequence 8, Appl1
30	1054	23.3	1079	4	US-08-484-159-8	Sequence 8, Appl1
31	1009.5	22.3	877	4	US-09-619-353-2	Sequence 2, Appl1
32	1002.5	22.2	843	4	US-09-361-631-1	Sequence 1, Appl1
33	999.5	22.1	863	4	US-09-619-353-14	Sequence 2, Appl1
34	974	21.5	863	4	US-09-619-353-14	Sequence 14, Appl1
35	938.5	20.7	1219	2	US-08-687-289A-6	Sequence 6, Appl1
36	839	18.5	835	4	US-09-619-353-7	Sequence 7, Appl1
37	788.5	17.4	851	4	US-09-619-353-12	Sequence 12, Appl1
38	787	17.4	856	4	US-09-619-353-8	Sequence 8, Appl1
39	781	17.3	854	4	US-09-619-353-10	Sequence 10, Appl1
40	756	16.7	669	4	US-09-361-631-7	Sequence 7, Appl1
41	683.5	15.1	1058	2	US-08-687-289A-5	Sequence 5, Appl1
42	672.5	14.9	872	3	US-08-337-797A-2	Sequence 2, Appl1
43	672.5	14.9	872	3	US-09-258-523-2	Sequence 2, Appl1
44	622	13.7	912	4	US-08-617-785-2	Sequence 2, Appl1
45	622	13.7	912	4	US-09-641-318-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-134-513-2
; Sequence 2, Application US/09134513
; Patent No. 6210964
GENERAL INFORMATION:
APPLICANT: Brown, Edward M.
APPLICANT: Diaz, Ruben
APPLICANT: Bai, Mei
APPLICANT: Quinn, Stephen J.
TITLE OF INVENTION: The Avian Extracellular Calcium-Sensing
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Vinson & Elkins L.L.P.
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20004-1008
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,513
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sanzo, Michael A.
REGISTRATION NUMBER: 36,912
REFERENCE/DOCKET NUMBER: BRI331/13003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)639-6585
TELEFAX: (202)639-6604
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1059 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-134-513-2

Query Match 24.1%; Score 1089.5; DB 4; Length 1059;
Best Local Similarity 30.4%; Pred. No. 1.7e-95;
Matches 266; Conservative 157; Mismatches 350; Indels 103; Gaps 23;
Qy 26 SQLRMKGDVYLGLPLG---EAEAGLRSTRPSSPVCTRFSSNGLMLAMKMAVEE 82

```

Db      23  NQRAOKKDDIILGFIHFVGAADKODIKS--RPESEVCIRYNERGRHMLQAMFAIEE 80
QY      83  INNSDILPGRLGYDEEDTCSEPVYAMKPSLMFLA--KASRDIAACANTQOQPRVLA 140
Db      81  INNSPMLPNTLGRITFDICNTVSKALEATLTFVAOKIDSLDNCSEHLPSTIA 140
QY      141  VIGHSESLAMVYTKRFSFELMPQVSYGASMLLSARETSPFRVPSDVLQALAE 200
Db      141  VVGTATGCVSTAVANLGLFYIPQVSYASSRLLSNKQFSLPTLIPNDEHQATAMADI 200
QY      201  LQEGMNVVALGSDDEYGRGLSTFSALAAAGICIAHEGLVPLPRADSRIGKVDYL 260
Db      201  IEFRRMNVGTIAADDDGRGIEKFRREAEERDIDFSELI---SQYDEEELQGVV 256
QY      261  HQVQSSVOVLLFASVAAAHAFNYSSISRLSPKVVAVASAMLTSDVLMGLPMAQGT 320
Db      257  EVIONSTRARIVYVSSSGDEPLKEIYRKNITGKIMLASAMASSSLIAMPFEFFVYGS 316
QY      321  VLGFLQGAQLHEFPQVY-KTHLALANDPACALGE----- 356
Db      317  TIGFALAKAGIIPGFREFQKVPKKSANNGFAKEWEETFCYLLPSEKNSPASAFHKA 376
QY      357  REOGL-----BEDVVGQRCPQDCDITLQNVAGLNNHOTSEVYAAVSYA 401
Db      377  HEEGLAGNGTAARPPCTGENTSVETPYMDTHLR-----ISYVYLAVYSIA 427
QY      402  QALNHTLOCNA-----SGCPADPVKPMOLLENMYLTFHVY-GLPFRDSSGNVME 453
Db      428  HALDITCTPGKGLFTNGSCADIKKYAMOVYKHLRLNTSMNGEYVDDEGDLVGN 487
QY      454  YDLKLM--VMQSVPLHDVGRFN-GSLRTERL---KIRWMTSDNQKFEVSRCSQO 504
Db      488  YSIINMLSPBDGSVY-FEEVGHYVYAKKGERLFINENKILMSGFSEKVEFPNCSRDCL 546
QY      505  EGOVRR-VKGFHSCCYDVCDEASRYRNPDIACITPCGODEMSPERKTRPFRRSRLA 563
Db      547  PGRKGIIEGPTCCFECVDCPDEYSDSETASACDKCPREYWSNENNTSCIPQOIELS 606
QY      564  WGEPAVILLLLLSLALGLVLAALGLFVHNDSPVQASGGLA---CFGVLCGLVLC- 619
Db      607  WTEPFGIALTFVYVGLFTSVGLVFTKFNTPVYKATNELSLVLLFSILC---CFS 662
QY      620  SVLFPQPSARCLAOQPLSHLPITGLSTLFLQAAIEY--ESELPLSNADRLSGCLR 677
Db      663  SLEFFIGPQWMTCLRLRPAFISFVLCISCLVKNRVLVFEAKIPTS-----LH 714
QY      678  GPW-----AMLVLLAMVLEVALCTWYLVAPPEVYDMMHLPTEAL-VHCRTSWSFG 731
Db      715  RKMWGLMQLPLVFLCTFVQIVICYIMLYTAPSSYRN-HELEDEILITCHEGSLMALG 773
QY      732  LAHATNATLAFICFLGTFVYRSGPCYNRANGLTPEAMLAFTVWSEFPLLANVQVLRP 791
Db      774  FLIGYTCUAAICFEFFAFKSKRLPENNEAKFITFSMLIFITVWISFLPAASVYGRFVS 833
QY      792  AVQMGALLCYLGLAFLHPRCYLMLRQPLNTP 827
Db      834  AVEVIAIILAAISFGILACIFENKVIYIILFKPSRNTIE 869

```

RESULT 2

```

; Sequence 7, Application US/08485588
; Patent No. 5688938
; GENERAL INFORMATION:
; APPLICANT: Edward M. Brown
; APPLICANT: Steven C. Hebert
; APPLICANT: Forrest H. Fuller
; APPLICANT: James E. Garrett, Jr.
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,588
; FILING DATE: 7 June, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below: 9
; APPLICATION NUMBER: 08/353,784
; FILING DATE: 9 December, 1994
; APPLICATION NUMBER: PCT/US/94/12117
; FILING DATE: 21 October, 1994
; APPLICATION NUMBER: U.S. 08/292,827
; FILING DATE: 23 August, 1994
; APPLICATION NUMBER: U.S. 08/141,248
; FILING DATE: 22 October, 1993
; APPLICATION NUMBER: U.S. 08/009,389
; FILING DATE: 23 February, 1993
; APPLICATION NUMBER: U.S. 08/017,127
; FILING DATE: 12 February, 1993
; APPLICATION NUMBER: U.S. 07/934,161
; FILING DATE: 21 August, 1992
; APPLICATION NUMBER: U.S. 07/834,044
; FILING DATE: 11 February, 1992
; APPLICATION NUMBER: U.S. 07/749,451
; FILING DATE: 23 August, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Hebert, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 213/005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1078 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-588-7

Query Match 23.6%; Score 1068.5; DB 1; Length 1078;
Best Local Similarity 30.3%; Pred. No. 1.8e-93;
Matches 269; Conservative 156; Mismatches 366; Indels 97; Gaps 26;

QY 12 WALL-----HPGTAPLCLSQGLRMKGDYVIGLFPPLG---EAEAGARSRTRPSSPYCTR 64
Db 8 WVLLATWHTSAYGP---DQRAOKKGDIIILGGLPIHFVGAADKODIKS--RPESEVCIR 62
QY 65 FSSNGILMALAMKVAEINNNKSDLLGLRLGYDLPFTCSEPVYAMKPSLMFLA--RAGS 122
Db 63 YNFRGFRMLQAMTFAIEINSSPALNPLTLGRIPTCNTVSKALEATLSFVAQONKIDS 122
QY 123 RDIACNYTOYOQPRVAVIGPSSSELMVYTKRFSFELMPQVSYGASMLLSARETFPS 182
Db 123 LINDPEFNCSEHIPSTIAVVGATGSGVSTAVANLGLFYIPQVSYASSRLLSNKQFVS 182
QY 183 FFRVPSDROVQTRAABELQEGMNVVALGSDDEYGRGLSTFSALAAAGICIAHEGL 242
Db 183 FLRTIPNDEHQATAMADIIEYFRMNVGTIAADDDGRGIEKFRREAEERDIDFSELI 242

```

```

OY 243 VPLPRADSRIGKVDVHLQVNOSSVOYVLLFASVAAHALFNYSISSRLSPKVVASAA 302
DB 243 I-----SOYSEDEEIOHVEVIONSTAKVIVVSSGDBLEPLIKEIVRRNITKIWLASBA 298
OY 303 WLTSDLVYKLPQMAOMGVYVLFQAGAOULHEPQYV-KTHLALADPAFCALSAGERE--- 358
DB 299 WASSSLIAMPQYFHVYVGGTIGFALKAGQIPGRREFLKVKHPKSVHNGAKFEWETENC 358
OY 359 -----OGIEEDVVGORCPQCD-----CITLQNVSA-----GLNHQ- 389
DB 359 HLOEGAKGPLVDYTLRGHEES--GDRFNSNSTARPLCTGDNENISSVETPYIDYTHLRI 416
OY 390 TFSVYAAVYVAQALHNTLQ-----NASCPADQVVKPMOLLENNYNTLF-HVGL 440
DB 417 STNVILAVYSIAHALODITYCLPGRGLFNGS-CADIKKVEAMOVULKHLRHLNFTNNMGE 475
OY 441 PLRFDSSGVNDMEYDLKLM--VMQSVPRLDHVGFRN-GSLRTERL-----KIRWHTSD 491
DB 476 QVTFDECGDLVGNYSIIMWHLSPEDGSIV-FKEVGYNNYAKGERLFINEEKILMSGFS 534
OY 492 NOKPVSRCROQEGQVR-VKGFHSCCYDCVCEAGSTRONPDIACTFCGODEWSPER 550
DB 535 REVPSNCSRDCLACTRGKIIIEGPTCCFCEVCECPDEGSEDTDSACNKCDDPMNSMEN 594
OY 551 STRCFRRSRFLWGPVAVILLLLSTLALGLVLAALGLFVHHRDSPVQASGFLA--- 607
DB 595 HTSCIAKELEFLSTWEPGIALTLPAVLGIFLTAFLVGLFIFRNPPIYKATNRELSTYL 654
OY 608 CGLVCLGLVCL-SVLEPPGQSPARCIAOPLSHPLTGLSTLFLQAEIFV--ESEL 664
DB 655 LFLSLC-----CRSSLEFGEFQDWTCLRQAPAGISFLVLCISLIVKNNRLVLEAKI 710
OY 665 PLS-----WADRLSCGLGPMANLYVLLAMVEVALCTWYLVAFPEVYTDHMLPTEL 719
DB 711 PTFSEHRKWMGLNQ-----FLLVFCTEMQIVICVIMLTAPSSYRNOLDEILF 762
OY 720 VHCRTSRWVSFGLAHATNATLAFCLFGLTFLVRSQPCYNRARGTLFAMLFITWSPV 779
DB 763 ITCHGSLMALAGFLGYCYCLAALCFPFAPKSRKLPENNEKFTFSFLIFFIYWI 822
OY 760 PLIANVOYVLRPAVOMGALLLCVLTGILAFHILPRCYLLMRQPLNTP 827
DB 823 PAVASTYKFSVAVEVIAIILAAFGLLACIFENKIYIILFKPSRMTIE 870

RESULT 3
US-08-484-565-7
; Sequence 7, Application US/08484565
; Patent No. 5763569
; GENERAL INFORMATION:
; APPLICANT: Edward M. Brown
; APPLICANT: Steven C. Hebert
; APPLICANT: James E. Garrett, Jr.
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESS: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,565

```

```

; FILING DATE: 7 June, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below: 9
; APPLICATION NUMBER: 08/353,784
; FILING DATE: 9 December, 1994
; APPLICATION NUMBER: PCT/US/94/12117
; FILING DATE: 21 October, 1994
; APPLICATION NUMBER: U.S. 08/292,827
; FILING DATE: 23 August, 1994
; APPLICATION NUMBER: U.S. 08/141,248
; FILING DATE: 22 October, 1993
; APPLICATION NUMBER: U.S. 08/009,389
; FILING DATE: 23 February, 1993
; APPLICATION NUMBER: U.S. 08/017,127
; FILING DATE: 12 February, 1993
; APPLICATION NUMBER: U.S. 07/934,161
; APPLICATION NUMBER: U.S. 07/834,044
; FILING DATE: 11 February, 1992
; APPLICATION NUMBER: U.S. 07/749,451
; FILING DATE: 23 August, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 213/006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1078 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-484-565-7

Query Match 23.6%; Score 1068.5; DB 1; Length 1078;
Best Local Similarity 30.3%; Pred. No. 18e-93;
Matches 269; Conservative 156; Mismatches 366; Indels 97; Gaps 26;

OY 12 WALL-----HGTGAPCLCSQOLRMKG DYVGLFPG---EAEAGLRTRTPSSPVCR 64
DB 8 WLLALTWHTSAVGP---DQRAOKGDITLGLFPHFGVAAKDDOKS--RPESECR 62
OY 65 FSSNGILMALAMKMAVEEINNKSDLLPGLRLGYDLFTCSEPVYAMKPSIMFLA--KAGS 122
DB 63 YNRGFRWLQAMIFALEEINSSPALPLNLTGRIPTDCTNFSKALEATLSFVAQNKIDS 122
OY 123 RQIAVYNTQYQPRVLAIVGPHSSSLANVTQKPFSEFLMPQVSGASWELLSARTFPS 182
DB 123 LNLDEFCNCSSEHPTIAVAVGATGSGVSTAVANLLGLFYIPQVSVASSRLLSNKQFES 182
OY 183 FFRVPSDRVQTLAAELLQEEFGMMVVAALGSDDEYRGOLSFSAALARGICIAHEGL 242
DB 183 FLRTIPNDHQATAMADILEYFRMMVVGITAADDDYGRGICIKFRBAEREOICIDFSL 242
OY 243 VPLPRADSRIGKVDVHLQVNOSSVOYVLLFASVAAHALFNYSISSRLSPKVVASAA 302
DB 243 I-----SOYSEDEEIOHVEVIONSTAKVIVVSSGDBLEPLIKEIVRRNITKIWLASBA 298
OY 303 WLTSDLVYKLPQMAOMGVYVLFQAGAOULHEPQYV-KTHLALADPAFCALSAGERE--- 358
DB 299 WASSSLIAMPQYFHVYVGGTIGFALKAGQIPGRREFLKVKHPKSVHNGAKFEWETENC 358
OY 359 -----OGIEEDVVGORCPQCD-----CITLQNVSA-----GLNHQ- 389
DB 359 HLOEGAKGPLVDYTLRGHEES--GDRFNSNSTARPLCTGDNENISSVETPYIDYTHLRI 416
OY 390 TFSVYAAVYVAQALHNTLQ-----NASCPADQVVKPMOLLENNYNTLF-HVGL 440

```

```

Db 417 SYNYTLAVYSTAHLADITYTCLPRGRLFTNGS-CADIKKEAVOMVLAHLNLTNNNGE 475
QY 441 PLRFDSSGNVMEYDLKLM---VMQSVPRLDHYGRFN-GSLRTERL-----KIRWHTSD 491
Db 476 QVTFDECDGLVGNYSIINMHLSPEDGSIV-EKEGYVNVAKKGERLFINDEKILMGSFS 534
QY 492 NQKPVSRCSRCQCGQVRR-VKGFHSCCYDCVCEAGSYRONPDIDACTCGGDEWSPER 550
Db 535 REVPFNSCRDCLAGTRGILIEGPTCCFECVCECPDEYSEDTASACNKCPCDFWSENEN 594
QY 551 STCFRRSRRLFAMGEPAVILLLLLSLALGLVLAALGLFVHHDSPLVYASGGPLA--- 607
Db 595 HTSCIAKEIEFLSWTEPRGIALTLFAYLGITLFAVLGFTKFRTPYVAKKNELSYLL 654
QY 608 CFGVLVGLVCL-SVLLFPQSPARCLAOPLSHLPYGLSLFLFOAIEFY--ESEL 664
Db 655 LFSILC-----CFSSSLFPIGEPQMTCLRPAPGISFVLCISCLVKNVLLVFEAKI 710
QY 665 PLS-----WADRLSGCLRGPMALVYLLAMVEVALCTWYLVAPPEVYDTMHLPTAL 719
Db 711 PTFHRRKMWGILNQ-----FLVFLCTFMQIVICIMLYTAPSSSYRQDEDEITF 762
QY 720 VHCRTSRWVSFGILAHATNATLAFICFTPLVRSOPGCYNRARGTLFAMLAFTWVSFV 779
Db 763 ITCHEGSLMAGFLIGTCLLAICFPFAFKSRKLPERENAKTITTSMLIFFTWISFI 822
QY 780 PLANVOVLRPAVOMGALLCVLGLLAFLHLPYCYLLMRQPGINTPE 827
Db 823 PAVASTGKFEVSAVEVIALIILASFGILACIFNKYIILIFKPSNHTIE 870

```

RESULT 4

US-08-480-751-7
Sequence 7, Application US/08480751

Patent No. 5858684

GENERAL INFORMATION:

APPLICANT: Edward F. Nemeth
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,751
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993

```

APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-751-7

```

Query Match

23.6%; Score 1068.5; DB 2; Length 1078;

Best Local Similarity 30.3%; Pred. No. 1.8e-93;

Matches 269; Conservative 156; Mismatches 366; Indels 97; Gaps 26;

```

QY 12 WALL-----HGTGAPLCSOOLRMRKGDYVIGLPLG---EAEAGRSKRSPSPVCTR 64
Db 8 WVLLALTWHTSAVGP---DQRAQRKGDIIIGLPLPHGVAAKQDIDKS--RPSVVCIR 62
QY 65 FSSNGILMALAMKAAVEININKSDLLPLRGLRYLPTDCSEPVYAMKPSIMFLA--RAGS 122
Db 63 YNFRGFRMLQMTAIEINISPPALNLTGYITFDICNTVSALEATISVAQNKIDS 122
QY 123 RDIAYCNTGYOYOPRLAVIGPSSSELMYTGKFFSEFLMPQVSGAMSELLSARETFS 182
Db 123 LINDFCNSHIPSTIAVAVGATGSGVSTAVANLLGLFYIPQVSYASSRILSKNFKS 182
QY 183 FFRIVPSDRVQTLAAELDFEGNNWYAAAGSDDEYGRGICSTISALAAAGCITIAEGL 242
Db 183 FLRTIPNDEQATMADIIIEFRNNWGTIAADYGRPGIEKREBAERDIDISEL 242
QY 243 VPLRADSRIGKYODVILHOVNOSSVOVLLFASVHAHALFNYSISRLSPKYVASEA 302
Db 243 I-----SQSDDEELQHYVEVYQNSTAYIVFSSGPDLEPLIKETIVRNITGKIWLASA 298
QY 303 WLTSDLVWGLPGMAQMGTVLGLQGAOLHEFPQYV-KTHLALATDPAFCASAGERE-- 358
Db 299 WASSSLIAMPQYFHVVGITGFALKAGQIPQFRFLKRVKSVHNGFAKEFEETFCNC 358
QY 359 -----QGLEDDVYGRCPQCD-----CITLQVNSA-----GLNHHQ- 389
Db 359 HLOGAKPLPVDTFLGLGHBS--GDRFNSSTAFRLCAGDENISVETPYIDYTHIRI 416
QY 390 TFSYAAVYSVAOALHNTLOC-----NASGCPADDPYKPMOLNMYNLTF--HYGGL 440
Db 417 SYNYTLAVYSTAHLADITYTCLPRGRLFTNGS-CADIKKEAVOMVLAHLNLTNNNGE 475
QY 441 PLRFDSSGNVMEYDLKLM---VMQSVPRLDHYGRFN-GSLRTERL-----KIRWHTSD 491
Db 476 QVTFDECDGLVGNYSIINMHLSPEDGSIV-EKEGYVNVAKKGERLFINDEKILMGSFS 534
QY 492 NQKPVSRCSRCQCGQVRR-VKGFHSCCYDCVCEAGSYRONPDIDACTCGGDEWSPER 550
Db 535 REVPFNSCRDCLAGTRGILIEGPTCCFECVCECPDEYSEDTASACNKCPCDFWSENEN 594
QY 551 STCFRRSRRLFAMGEPAVILLLLLSLALGLVLAALGLFVHHDSPLVYASGGPLA--- 607
Db 595 HTSCIAKEIEFLSWTEPRGIALTLFAYLGITLFAVLGFTKFRTPYVAKKNELSYLL 654

```


QY 608 CEGIVCLVCL-SVLEFPQSPARCLAQPLSHLPLTGLSTFLQAEIV--ESEL 664
 Db 655 LPSLCLC-----CFSSLEFTEGPODWTCKLRQPAFISVLCISCLVNTKRNVLVFEAKI 710
 QY 665 PLS-----WADRLSGCLGPNAMLVLLAMVEVALCTWYLVAPPEVVTDMHMLPTAL 719
 Db 711 PTFPRKRWGINTLQ-----FLIVFLCTFMQIVICIMYLTAPSSRYRNOLEDELIIF 762
 QY 720 VHCRTSRVSVFGLAHATNATIAFLCTGTFIVRSOPGCVNARAGITFAMLYFTWVSFV 779
 Db 763 ITCHGSLMALGFLIGYICLAAICFFPAFKSRKLPENFNKAKFTFEMLIFFIWMISFI 822
 QY 760 PLANVOYVLRPAVOMGALLICVGIILAAFLHPLRCYILMRQPLNTPPE 827
 Db 823 PAYASTYKRFVSAVEVIAIILASFGLLACITFNKITYIILFKPSRNTIE 870
 RESULT 5
 : Sequence 7, Application US/08943986
 : Patent No. 5962314
 : GENERAL INFORMATION:
 : APPLICANT: Edward M. Brown
 : APPLICANT: Steven C. Hebert
 : APPLICANT: James E. Garrett, Jr.
 : TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 : TITLE OF INVENTION: MOLECULES
 : NUMBER OF SEQUENCES: 20
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Lyon & Lyon
 : STREET: First Interstate World Center
 : STREET: Suite 4700
 : STREET: 633 West Fifth Street
 : CITY: Los Angeles
 : STATE: California
 : COUNTRY: USA
 : ZIP: 90071
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 3.5" Diskette; 1.44 Mb storage
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: FASTSEQ
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/943,986
 : FILING DATE: 03-OCT-1997
 : CLASSIFICATION: 530
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/484,565
 : FILING DATE: 7-June-1995
 : APPLICATION NUMBER: 08/353,784
 : FILING DATE: 9 December, 1994
 : APPLICATION NUMBER: PCT/US/94/12117
 : FILING DATE: 21 October, 1994
 : APPLICATION NUMBER: U.S. 08/292,827
 : FILING DATE: 23 August, 1994
 : APPLICATION NUMBER: U.S. 08/141,248
 : FILING DATE: 22 October, 1993
 : APPLICATION NUMBER: U.S. 08/009,389
 : FILING DATE: 23 February, 1993
 : APPLICATION NUMBER: U.S. 08/017,127
 : FILING DATE: 12 February, 1993
 : APPLICATION NUMBER: U.S. 07/934,161
 : FILING DATE: 21 August, 1992
 : APPLICATION NUMBER: U.S. 07/834,044
 : FILING DATE: 11 February, 1992
 : APPLICATION NUMBER: U.S. 07/749,451
 : FILING DATE: 23 August, 1991
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Hebert, Sheldon O.
 : REGISTRATION NUMBER: 38,179
 : REFERENCE/DOCKET NUMBER: 213/006
 : TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 : INFORMATION FOR SEQ ID NO: 7:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1078 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-08-943-986-7
 Query Match 23.6%; Score 1068.5; DB 2; Length 1078;
 Best Local Similarity 30.3%; Pred. No. 1,8e-93;
 Matches 269; Conservative 156; Mismatches 366; Indels 97; Gaps 26;
 QY 12 WALL-----HPGTGAPLCISQQLRMKGDVYLGLEPLG--EAENAGLSRFRSPSPVCTR 64
 Db 8 WVLLATLWHTSAYGP---DORAKKKDIIILGLPFPHGVAAKQDDLS--RRESVEICIR 62
 QY 65 FSSNGLMALAMKMAVEEINKKSDLLPGLRLGYDLFDTCSEFPVAMKRSMLFLA--KAGS 122
 Db 63 YNFKGFRLQAMITALEEINSPPALLPLTLGYRLFDTCNTVSKALEYLSFVAKNKIDS 122
 QY 123 RDIAYCNVTOYOPRVLAIVIGPHSSELMVTKFESFLMPQVSYGASMEILSAETPPS 182
 Db 123 LINDFECNSHIPSTIAVAGATGCVSTAVANLLGLFIYIQVSYASSRLSNKQPKS 182
 QY 183 FFRVPSDRVDTAAAEILOEFGNMVVALGSDDEYGRGSLITSALANGICTIAHEGL 242
 Db 183 FLRTIPNDHQATMADIEFEFRNMVGTIAADDYGRPGIEKEFEEREEEDICIDFSEL 242
 QY 243 VPLPADSRIGKQVDVHOVNOSVOYVLLFASVHAHALFNYSISRLSPKVVAVSEA 302
 Db 243 I-----SOYSDEEIOHVEVIONSTAKIVVSSGPDLEPLKEIVRNITGKMLASEA 298
 QY 303 WLTSDLVNGLPGMOMGTVLGFLQGAOLHEFPQYV--KTHLALATDPAFCSALGERE--- 358
 Db 299 WASSSLIAMPQYFHVYGGTIGFALKAGQIPGRFELKVVHPRKSVHNGFAEFMEYETNC 358
 QY 359 -----QGLEDDVYQRCPOCD-----CITLQVSA-----GLNHQ- 389
 Db 359 HLOEGAKGPLVDQFLRGHEBS--GDRPSNSSTAFRLPCLTDENSSVETPYIDYTHIRI 416
 QY 390 TFSYAAVYSVAOALHNTLOC-----NASGPAQDPVAPWOLLEMYNLTFF--HYGGL 440
 Db 417 SYNTYLAIVYSTAHALQDITYTLPGRGLETNGS--CADIKYVAMQVYKLRHLNFTNNNGE 475
 QY 441 PLRFDSSGNVDMEXDLKLM---VMQGSVRLHDVGRFN--GSLFRERT-----KIRWHTSD 491
 Db 476 QVTFDECGDLVGNYSIIMMHLSPEDGSIV--FREYGVYVYAKKGRFLFINBEKILMSGFS 534
 QY 492 NQKPRSRCSROQSQVNR--VKGRHSCICYDCEAGSYRONPDIDACTFGCODEMSER 550
 Db 535 REVPSPNSCRDLGTRGRIIEGEPTECFCEVCEPDGYSDETSAACNKPDPDFWSENEN 594
 QY 551 STCFRRSRFLANGEPVAVLLLLLSLALGLVLAALAEFLVHHRSPLVQASGGLA--- 607
 Db 595 HTSCIAKIEFLSTWTERPGIALTLFAVIGITLFAVGVIFKFRNTPIYKATFNELSL 654
 QY 608 CEGIVCLVCL-SVLEFPQSPARCLAQPLSHLPLTGLSTFLQAEIV--ESEL 664
 Db 655 LPSLCLC-----CFSSLEFTEGPODWTCKLRQPAFISVLCISCLVNTKRNVLVFEAKI 710
 QY 665 PLS-----WADRLSGCLGPNAMLVLLAMVEVALCTWYLVAPPEVVTDMHMLPTAL 719
 Db 711 PTFPRKRWGINTLQ-----FLIVFLCTFMQIVICIMYLTAPSSRYRNOLEDELIIF 762
 QY 720 VHCRTSRVSVFGLAHATNATIAFLCTGTFIVRSOPGCVNARAGITFAMLYFTWVSFV 779
 Db 763 ITCHGSLMALGFLIGYICLAAICFFPAFKSRKLPENFNKAKFTFEMLIFFIWMISFI 822
 QY 780 PLANVOYVLRPAVOMGALLICVGIILAAFLHPLRCYILMRQPLNTPPE 827

Db 823 PAVASTGKFSAVEVAIALLAASFGLLACJFENKIYIILLKPSRNTLE 870

RESULT 6

US-08-353-784-7

Sequence 7, Application US/08353784

Patent No. 6011068

GENERAL INFORMATION:

APPLICANT: Edward F. Nemeth, Edward M.

APPLICANT: Brown, Steven C. Hebert,

APPLICANT: Bradford C. Van Wageningen, Manuel

APPLICANT: F. Balandrin, Forrest H. Fuller,

APPLICANT: Eric G. Delmar, and Scott T. Moe

TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: First Interstate World Center

STREET: Suite 4700

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/353,784

FILING DATE: 9 December, 1994

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA: including application

PRIOR APPLICATION DATA: described below: 8

APPLICATION NUMBER: PCT/US/94/12117

FILING DATE: 21 October, 1994

APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994

APPLICATION NUMBER: U.S. 08/141,248

FILING DATE: 22 October, 1993

APPLICATION NUMBER: U.S. 08/009,389

FILING DATE: 23 February, 1993

APPLICATION NUMBER: U.S. 08/017,127

FILING DATE: 12 February, 1993

APPLICATION NUMBER: U.S. 07/934,161

FILING DATE: 21 August, 1992

APPLICATION NUMBER: U.S. 07/834,044

FILING DATE: 11 February, 1992

APPLICATION NUMBER: U.S. 07/749,451

FILING DATE: 23 August, 1991

ATTORNEY/AGENT INFORMATION:

NAME: Hebert, Sheldon O.

REGISTRATION NUMBER: 38,179

REFERENCE/DOCKET NUMBER: 209/069

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1078 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-353-784-7

Query Match 23.6%; Score 1068.5; DB 3; Length 1078;

Best Local Similarity 30.3%; Pred. No. 1,8e-93;

Matches 265; Conservative 156; Mismatches 366; Indels 97; Gaps 26;

QY 12 WALL-----HPGTAPLCSQOLRNKGDYVIGLFPIC---EAEAGLRSTRPSSPCSTR 64

Db 8 WVLLATWHTSAYGP---DQRAQKGDIIILGLPIPHFVAAKQDQDKS--RPESVACIR 62

QY 65 FSSNGLLMALMKVAEINKKSDLLPGIRGYPDPTCSPPVAMPSLMFLA--KAGS 122

Db 63 YNFRGFRMLQAMIPALIEINSSPALNLTGLYKIPDTCNVSKALNATLSFVQNMKIDS 122

QY 123 RDIAYCNYTOYOPRYAVIGPSSSELMWYKFFSEFLMPOYSYGASMELSARETPS 182

Db 123 LNDPEFCNSHIPSTIAVAGATGSGVSTAVANLGLFYIPQVSYASSRLSKNFKS 182

QY 183 FETVPDROVLTAAELDFEGNMWVAALGSDDEYRQGLSTPSALAAAGICIAHEGL 242

Db 183 FLRTIPNDHQATAMADIIIEYFRNMWGTIAADDYGRPGIEKREAREEDICIDSESEL 242

QY 243 VPLPRADSRKGVQVDLHQVNOSSQVVLFLASVHAHLFNYSISRLSPKVVASEA 302

Db 243 I---SOYSDDEETOHVEYIQNSTAVYVSSGPDLEPIKEIVRNITGKIWLASEA 298

QY 303 WLTSDLVWGLPGMAQMGTVLFGQAQLHEFPQYV-KTHLATATDPAFCSALGERE--- 358

Db 299 WASSSLIAMPOYFHVYGGTIGFALKAGQIIPGREFLAKVHRKRVHNGFAKEFWDETFNC 358

QY 359 -----OGLEEDVVGORPOCD-----CITLQNVSA-----GLNHQ- 389

Db 359 HLAGAKGPLEVDTPFLRGHEES--GDRFSNSTAFRLPCTGDEMISSEVPEYIDYHLRI 416

QY 390 TFSYAAVYSAOALNHTLQC-----NAGCPADDPYKPMQLLNMNLT/-HNGGL 440

Db 417 SYNVYLAIVYSTAHALODIYTCPLPGRGLFTNGS--CADIKRYEAMOVVLAHLRLNTNNGE 475

QY 441 PLRFSSGNVMEYDLKLM--VWQGSVPRLDHGRN-GLRERLT-----KIRHTSD 491

Db 476 QVTFDECGDLVGNKSIINMHLSPEDGSIV-EKEGYINVAKKEGRFLINEKILMNGFS 534

QY 492 NQRFVSCRSQOCQGYR-VKGFHSCCYDCVDCAGSYRONPDIACTCGODENSEPER 550

Db 535 REVFVSCSRCLAGTRKGIIEGEPCEPDEYSDPDSACNKCPCPDFWGNEN 594

QY 551 STCFRRSRRLAMGEPAVILLILLLSALGLVLAALGLFVHNDSPLYONASGSLA--- 607

Db 595 HTSCIAKEIEFLSWTEPPGIALTLFANVIGITLFAVIGVLFKFNNTIYKATNRELYL 654

QY 608 CFGLVCLGLVCL-SVLLPFGQSPSPARCLADQPLSLPLTGCLSTYFLQAAEIFV--ESEL 664

Db 655 LFSLLC-----CFSSSLFFIGEPQDMTCRLRPAGISFVLCISCLIVKTRNVLLVFEAKI 710

QY 665 PLS-----WADRLSGCLRGPMALVYLLAMLEVVALCTWYLVAPPEVYTDHMLPTREAL 719

Db 711 PTFHRRKMWGINTQ-----FLVFLCTFMQIVICVIMLYTAPSPSYRNOLEDEITF 762

QY 720 VHCRTSWSVSGFLAHATNATLAFICFLGTFLVRSQPCGVNARGLTAMLAYFTWVSFV 779

Db 763 ITCHEGSLMALGFLIGTCLLAICFFPAFERSKRLPENFENAKITTSMLFTWLSFI 822

QY 780 PILANQVVLPAVQMGALLCYGLIILAFTIPRCYLLMRPGINTEP 827

Db 823 PAVASTGKFSAVEVAIALLAASFGLLACJFENKIYIILLKPSRNTLE 870

RESULT 7

US-08-484-719B-7

Sequence 7, Application US/08484719B

Patent No. 6031003

GENERAL INFORMATION:

APPLICANT: Edward F. Nemeth, Edward M.

APPLICANT: Brown, Steven C. Hebert,

APPLICANT: Bradford C. Van Wageningen,

APPLICANT: Manuel F. Balandrin,

APPLICANT: Forrest H. Fuller, Eric G.

TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE

```

1 TITLE OF INVENTION: MOLECULES
2 NUMBER OF SEQUENCES: 20
3 CORRESPONDENCE ADDRESS:
4 ADDRESSEE: Lyon & Lyon
5 STREET: First Interstate World Center
6 STREET: Suite 4700
7 STREET: 633 West Fifth Street
8 CITY: Los Angeles
9 STATE: California
10 COUNTRY: USA
11 ZIP: 90071
12
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
15 COMPUTER: IBM PC compatible
16 OPERATING SYSTEM: MS Word
17 SOFTWARE: FastSeq for Windows Version 3.0
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER: US/08/484,719B
20 FILING DATE: 7 June, 1995
21 CLASSIFICATION: 514
22
23 PRIOR APPLICATION DATA:
24 APPLICATION NUMBER: 08/353,784
25 FILING DATE: 9 December, 1994
26 APPLICATION NUMBER: PCT/US/94/12117
27 FILING DATE: 21 October, 1994
28 APPLICATION NUMBER: U.S. 08/292,827
29 FILING DATE: 23 August, 1994
30 APPLICATION NUMBER: U.S. 08/141,248
31 FILING DATE: 22 October, 1993
32 APPLICATION NUMBER: U.S. 08/009,389
33 FILING DATE: 23 February, 1993
34 APPLICATION NUMBER: U.S. 08/017,127
35 FILING DATE: 12 February, 1993
36 APPLICATION NUMBER: U.S. 07/934,161
37 FILING DATE: 21 August, 1992
38 APPLICATION NUMBER: U.S. 07/834,044
39 FILING DATE: 11 February, 1992
40 APPLICATION NUMBER: U.S. 07/749,451
41 FILING DATE: 23 August, 1991
42
43 ATTORNEY/AGENT INFORMATION:
44 NAME: Douglas C. Murdock
45 REGISTRATION NUMBER: 37,549
46 REFERENCE/DOCKET NUMBER: 213/007
47 TELECOMMUNICATION INFORMATION:
48 TELEPHONE: (213) 489-1600
49 TELEFAX: (213) 955-0440
50 TELEX: 67-3510
51
52 INFORMATION FOR SEQ ID NO: 7:
53 SEQUENCE CHARACTERISTICS:
54 LENGTH: 1078 amino acids
55 TYPE: amino acid
56 STRANDEDNESS: single
57 TOPOLOGY: linear
58 MOLECULE TYPE: protein
59
60 US-08-484-719B-7
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
90
```

```

Db      183  FLRTIPNDEHQAATAADILIEFFRNWVGTTIAADDDYGRPGLEKREBAERDICIOPSEL 242
OY      243  VLPFPADSRSLRGKVODVLYHQVNOSSVOYVLLFASVHAHAALFNSSISSRPKWVWASEA 3020
Db      243  I-----SQSDEBEIDHVEYVIONSTAKYIVFSSGPDLEPLIKYIVARNITGKIWLASEA 298
OY      303  WITSDLYNGLEPGMAOMGTVLGFLORGAOLHEFPQYV- KTHLATATDPAFCALSAGERE-- 358
Db      299  WASSSLIAMPQYFHVHVGSTIGFALKACQIPGFRBEFLKVRHRSVHNGFAKFEWEETFC 358
OY      359  -----GSELEDVYGORCPQD-----CITLONVSA-----GLNHQ- 389
Db      359  HLQEGAKPRLPDVDFLRGHEES--GDRESNSTAFRPLCTGDEMISSVEPTYIDYTHLRI 416
OY      390  TESVYAAVYSVAOALHNLTLOC-----NASGCPADPYKPMOLNMTNLTFF-HVGG 440
Db      417  SINYVLAYYSIAHALQDITYLTPRGLETNGS -CADIKRYEAMQVYLKHLHLNFTNMGE 475
OY      441  PLREDSGNVDMEXDULKIM--VMQGSVPLRLHDYGRFN-GLALTERL-----KIRWHTSD 491
Db      476  QVTPECGDLVGNYSIIMWHLSPEDGSIV -PKEVGYNYVAKKERLFIINEKILMGFS 534
OY      492  NOKPYSRCSROCOBQVNR-VKGHSCCYDCVDCSEASYRONPPDICTCGODEMSPER 550
Db      535  REVPPSNCSROCLAGTRGIIIEGEPCTCFEEVCECPDEYSEDPTDASACNCPDPFMSNEN 594
OY      551  STRCRRRSRFLWAGEPRAVLLLLLLLSLALGLVLAALGFLVHHSDPLVQASG6PLA--- 607
Db      555  HTSCIAKIEFLSTWTEPGILTLFAYVIGTLFVAVGVFKERNYIVATNREISYLL 654
OY      608  CFGVLCGLVCL-VLLEPPGQSPARCLAQOPLSHLPLTGLSTLFLQAAEIVY--ESEL 664
Db      655  LFSLIC-----CFSSLFPIGBPDWMTCLRQPARGISFVLCISCIILKTRNVLVPEAKI 710
OY      665  PLS-----WADRLSCLRGPRAMLVYLLAMEVALCTWLYVAFPREVYTDHMLPREAL 719
Db      711  PFSFRKRWGMLQ-----FLVLECFPMQIVICVIMLYTAPSSSYRNOLEDEIIF 762
OY      720  VHCFTRSVWSFGLAHATNATLAFICEFTGTFVRSOPGCYNRARGLFPAMLAFTWVSFV 779
Db      763  ITCHEGSLMALGELIGYICLLAICFPFAFRSKRLPENFNKAITTSMILPFIWISFI 822
OY      760  PLIANVQVLRPAVOMGALLICVLGILIAAFHLPRCYLLMRPGINTPE 827
Db      823  PAVASTYGFVSAVEVIAIILASGILLACIFENKIYIILFRPSNITIE 870

RESULT 8
US-08-484-159-7
: Sequence 7, Application US/08484159
: Patent No. 6313146
: GENERAL INFORMATION:
: APPLICANT: Bradford C. Van Wagenen
: APPLICANT: Manuel F. Balandrin
: APPLICANT: Eric G. Del Mar
: APPLICANT: Edward F. Nemeth
: TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
: TITLE OF INVENTION: MOLECULES
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: First Interstate World Center
: STREET: Suite 4700
: STREET: 633 West Fifth Street
: CITY: Los Angeles
: STATE: California
: COUNTRY: USA
: ZIP: 90071
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,159
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 214/101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-159-7

Query Match      23.6%  Score 1068.5  DB 4:  Length 1078:
Best Local Similarity 30.3%  Pred. No. 1.0e-93;
Matches 269; Conservative 156; Mismatches 366; Indels 97; Gaps 26;
QY 12 WALI-----HPGTGAPLCLISQQLRMKGDYVLGGLPLG---EAEAGLSRTRPSPVCTR 64
DB 8 WVLLALWHTSAYGR---DQRAQKKGDIIILGGLPIHGVAAKODLKS--RPSVEGIR 62
QY 65 FSSNGLLMALAKMAVEINNKSLDPLGLRIGYDLFTGCEPVVAMKPSLMFLA--KAGS 122
DB 63 YNFGFEMLOAMTFALIEINSSPALLPLNLGRIPIFTCNTVSKALNTLSFVQNKIDS 122
QY 123 RDIAYCNYTOYORVAVIAGHSELAMVYGKFFSEFLMPOVSYGASMSLLSRETPPS 182
DB 123 LNDDEFNCSEHIPSTIAVVGATGSGVSTAVANILGLFTIPQVSYASSRLSKNPKS 182
QY 183 FFRTPSDRYVLAALAEILQFEGNNWVAALGSDDEYGRGLSIFSALAAAGICIAEGL 242
DB 183 FLRTIPDEHQATAMADIIIEFRMNWGTIAADDDYGRPGIEKREAEEDIDICIDSEL 242
QY 243 VPLPRADSRKGVQDVLHOVNOSSVOYVLLFASVHAHAHLENYSISSRLSPKYWASEA 302
DB 243 I-----SOYSDEEIOHVEVYIONSTAKVIVFSSGPDLEPLIKETIVRRNITGKIWLASEA 298
QY 303 WLTSDLMWGLPGMAOMGTVIGFLORGNOHHEFPQYV-KTHALATDPAFCALDEERE--- 358
DB 299 WASSSLIAMPOYFHVAGTIGFALKAGQIDGFEFLKRVHPRKSVHNGFAKEFEETPNC 358
QY 359 -----OGLIEDVVGORCPQCD-----CITLQWVA-----GLNHQ- 389
DB 359 HLDGKAGPLPVDITFLGHEBS--GDRFNSNSTAFRPLCTGDENISSVEPYIDYTHIRI 416

```

```

QY 390 TFSYAAVYSVAQALHNTLOC-----NASGCPADDPVPMOLNMYNLF-HVYGL 440
DB 417 SYNYLAVYSTAHLODIYICLPGRGLETNGS-CADIKKVAOMVYLHRLNLTNNMGE 475
QY 441 PLRFDSSGNVMEYDLKIM---VMQSVPLRHVDYGRN-GLRTERL-----KIRWHTSD 491
DB 476 QVTEDECDGLVGNYSIINWHLSPEDGSIV-FKEGYVNVYAKKGERLFINNEKILMSGFS 534
QY 492 NOKVVSQSCROCGOVR-R-KGFHSCCYDVCDEASYSRONPNPDICTFCGODEMSPER 550
DB 535 REVPFNSCROCLAGTRKGIIEGPTCCFECVCPDDESDETASACNKCDDPFWMSNEN 594
QY 551 STCRFRRSRRLWAGEPAVLLILLLSIALGLVLAALGLFVHHDSPLVYASGGPLA--- 607
DB 595 HTSCIAIEIEFLSWTEPFGIALTLFAVLGIFLAFVIGFVIFKFNPTIVKATNELSYLL 654
QY 608 CFGVCLGLVCL-SVLPFROPSPARCLAOQPLSHLPLTGLSTLFLQAAEIFY--ESEL 664
DB 655 LFSLLC---CFSSSLEFFIGEPQDMTCRLQRPAGISFVLICISILVKNRVLLVFEAKI 710
QY 665 PLS-----MADRLSGCLRGPMAMLVILLAMLEVALCTWYLVAPPEVYTDHMLPTEAL 719
DB 711 PTFHRRKMWGLNQ-----FLVFLCTFMQIVICVIMLYTAPSPSYRQGEJDEIIF 762
QY 720 VHCRTSMVSEGLAHATNATLAFICLTGTLVRSQPCYCNBARGLPAMLAFTTWSEV 779
DB 763 ITCHEGSLMALGFLIGYTCLLAACFFFAFKSRKLPENFNARKITFSMLFIETWISFI 822
QY 780 PLANNQVLRPAVOMGALLCVGIIAFLPRCICYLLMRQGLNTEP 827
DB 823 PAVASTYKREYSAVEVAILLAASFGLACIFENKIYITLEKPSRNTLE 870

RESULT 9
US-08-485-588-5
Sequence 5, Application US/08485588
Patent No. 5688938
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,588
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994

```

```

APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1085 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-588-5

```

```

Query Match      23.4%; Score 1060; DB 1; Length 1085;
Best Local Similarity 30.0%; Pred. No. 1.2e-92;
Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

```

```

QY 6 VGLSLMALHPGTGAPLILSLQGLKGVYVGLFPL--GEA--EENAGLRSTRSPSPVC 62
DB 10 LIAFSTWCSAIVP-----DRAQKKGDIILGGLPIHFGVAVKDQDLKS--RPSVVC 61
QY 63 TRFSNGILMALAMKAAVEINNKSDLPGLRIGDLPDTCSEPVYAMKPSIMFLA--KA 120
DB 62 IRYNFRGFRWLQAMIFAIKEINSSPALHPMLGRIEPTCNVTSALATLSFVANKI 121
QY 121 GSRDLAAYCNITQYQPRVAVIGPHSSSLAMVYTKFFSFLMPOVSYGSMELLSARETF 180
DB 122 DSLINDEFNCSEHPISTIAVVGAGSGISTAVANLGLFYIPQVSYASSRILSKNOF 181
QY 181 PSFRTVPDSRYOLFAALLOEFGMNWVAALGSDDEYRGOGISFSAALAAAGCIIAHE 240
DB 182 KSFLLRIPDEQATAMADIEYFRNMWGTIAADDYGRPGIEKFRERAEEDICIDIFS 241
QY 241 GLVPLPRADSRIRGVQVYLQHVNOSSVOVLLFASVHAHALFNYSISSRLSPKVVAS 300
DB 242 ELI-----SQYSDERKIQVVEVYIONSTAVIYVSSGPDLEPLIKETIRNITGRITLAS 297
QY 301 EAMLSIDLVMGLPGMAQMGTVLGTORGALHEFPQYV-KTHLALATDPAFCSAI----- 354
DB 298 EAMASSSLTAMPEYFHVGGTIGFGLKAGQIPGFFREFLOKVRHRSKSHNGFAKEWEETE 357
QY 355 -----GEREGL-----BEDVYGQCGPCDDCTTGN 380
DB 358 NCHLDGAGAPLPVDITFLNGHEGKARLSNSPTARPLCTGEBNISSTVETPMADYTHLR- 416
QY 381 VSAGLNHQTFSVYAAYSVAQALNHTLOC-----NASGCPADDPYKPMOLLENMTN 432
DB 417 -----ISYNYLAVVYSIAHALDITVICIPGRGLFTNGS--CADIKRYEAMOVYKLHLNH 467
QY 433 LTFHNG-GLPLRFDSSGNVDMEDYDKLW---VMQSVPLHLDVGRN--GSLTERL----- 483
DB 468 LNTFSMGQVTFDECGDLAGNYSIINWHLSPEDGSIV-FKEVGYNYVAKKGERLFIIND 526
QY 484 -KIRMTSDNOKPVSRCSQCOEGQVR--VKGFHSCCYCDOCEAGSYGNPDIDICTFC 541
DB 527 EKILNMGFRREYFSCSDICLAGIRKGIIEGPTCCFCEVCECPDEYSDETFDASACDKC 586
QY 542 GODEMSPERSTRCFRRRSRFLAMGEPAVILLILLISLALGLVLAALGLFVHHRDSPLVQA 601

```

```

DB 587 PDDFMSENHNTSCIAKEIFELSMTEPFGIALTLFENVLGFLAFLVGVIFKRNRPIVNA 646
QY 602 SCGPPLA---CFGLVCLGLVCL-SVLIFFGQSPBARCLAOQPSHLPLTCCSLTFLQAAE 657
DB 647 TUNELSYLLIFSLIC---CFSSSIFETEPDQWTCRLQOPAFGISFVLCISCIIVKTR 702
QY 658 IFV--ESELPLS-----MADRISGCLRGPMWLVVLAALVEVALCTWLVAFPEPVND 710
DB 703 VLVLEAKIPTSEHRKMGMLNQ-----FLVFLCTFMQIVICATMLNTAPSSYRN 754
QY 711 WHMLPTREAL-VHCRSRWVSFGIAHNTNATLAFELCLGFLVRSOPGCYNRRAGLFPAML 769
DB 755 -HEDEELIFITCHEGSLMALGFLIGYTCLLAICFEFAFKSKLPENENKAFTEFSML 813
QY 770 AFTITWVSFVPLLANQVYLRAVQMGALLCYLGLAFLHLPRLCYLMRGSLNTP 827
DB 814 IEFIVWISPIPAVASTYGRFVSVAEVIALAASFGLACIFPNKYIILFKRSRTIE 871

```

```

RESULT 10
US-08-484-565-5
; Sequence 5, Application US/08484565
; Patent No. 5763569
; GENERAL INFORMATION:
; APPLICANT: Edward M. Brown
; APPLICANT: Steven C. Hebert
; APPLICANT: James E. Garrett, Jr.
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; CITY: 633 West Fifth Street
; STATE: Los Angeles
; COUNTRY: California
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,565
; FILING DATE: 7 June, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: Including application
; PRIOR APPLICATION DATA: described below: 9
; APPLICATION NUMBER: 08/353,784
; FILING DATE: 9 December, 1994
; APPLICATION NUMBER: PCT/US/94/12117
; FILING DATE: 21 October, 1994
; APPLICATION NUMBER: U.S. 08/292,827
; FILING DATE: 23 August, 1994
; APPLICATION NUMBER: U.S. 08/141,248
; FILING DATE: 22 October, 1993
; APPLICATION NUMBER: U.S. 08/009,389
; FILING DATE: 23 February, 1993
; APPLICATION NUMBER: U.S. 08/017,127
; FILING DATE: 12 February, 1993
; APPLICATION NUMBER: U.S. 07/934,161
; FILING DATE: 21 August, 1992
; APPLICATION NUMBER: U.S. 07/834,044
; FILING DATE: 11 February, 1992
; APPLICATION NUMBER: U.S. 07/749,451
; FILING DATE: 23 August, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179

```

REFERENCE/DOCKET NUMBER: 213/006
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1085 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-480-565-5

Query Match 23.4%; Score 1060; DB 1; Length 1085;
 Best Local Similarity 30.0%; Pred. No. 1,2e-92;
 Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

QY 6 VLGLSLALHPGAPGLCSQQLRMKDDYVGLGLPL--GBA-EEAGLRKRTSPSSPVC 62
 DB 10 LLAFTWCTSAAYP-----DQRAQKGDIIIGLFIHFVAVAKDOLKS--REPSEYEC 61
 QY TRFSSGLLALAMKAVEEINNKSDILPLRLIGYDLPDTCSEPVVAMKPSIMFLA--KA 120
 DB 62 IRYNFRGRMTQAMTFAIEEINSSPALLPNNLTGTRIDTCNVSKALEATLSFVAONKI 121
 QY 121 GSRDIAAYCNTQYQRYAVIAGPSSSELAVYTKFESFPLMPQVSYGASMSLLSARETF 180
 DB 122 DSLNIDFRCNCSEHPISTAVVATGSGISTAVANMLGLFTIPQSVASSSSRLSNKNOF 181
 QY 181 PSFRTVYSDRVQVLAABELQERGMNVVAAAGSDDEYRGGLSFFSLAAARGCICNHE 240
 DB 182 KSFLETPNDHQATAMADIIIEYFRMNWVGTTAADDGRRGIEKFRREAEERDIDFS 241
 QY 241 GLVPLPRADSRIGKQVDVLHQNVSQVYVLLFASVAAHAFNYSSISRLSPVWVAS 300
 DB 242 ELI-----SQYDEEKIQVVEIONSTAKVIVPSSGPDLEPLKEIYRNTGIMWLAS 297
 QY 301 EAMLTSLVWGLPGMAQGYVGLQGAQLHEPQYV-KTHLALATDPAFCSAL----- 354
 DB 298 EAMASSSLIAMPDEYHVVGITIGFELKAGQIPGREFLOKVPKRKSVHNGFAKEFEWETF 357
 QY 355 -----GEREGL-----EEDVVGRCQPCOCITON 380
 DB 358 NCHLOEAKKGLPVDFTLRGHEGGARLSNSPTAFRPLCTGENSESVETPMDDTHLR- 416
 QY 381 VSAGLNHQTFSVYAAVYVAQALHNTLQC-----NASGCPAQDVPKPKWOLENNYN 432
 DB 417 -----ISYNVYAVIAHAIADIDITCIPGRGLFTNGS-CADIKKVEAMQVAKLHRH 467
 QY 433 LTFHYG-GLPLRFDSGGVNDMEYDKLM---VMGGSVRLHDVGRFN-GSLTERL---- 483
 DB 468 LNFYSNMGEQVTFDECGGLAGNYSIINWHLSPEDSIV-FKEVGYYNYAKGERLFTIND 526
 QY 484 -KIRHNTSDNCKPYRSGROCOEGQVR-VKGFHSCYDCVDCCEGSRONDDIACIFC 541
 DB 527 EKILMSGSRVPPSNCROCLAGRKGIIRGEPPCCFCEVCEPDEGSEDETASACCKC 586
 QY 542 GODEMSPERTRCFRRSRFLAMGEPVALLLLSLALGVLALGIFVHHRDSPLVOA 601
 DB 587 PDDFMSNNHNSCIATKEFLSWTEPEFIALTFRAVLGIFLAVLGIFINRNPPIYKA 646
 QY 602 SGGPLA---CGLYCLGLVCL-SVLLPQGPSPARCLAQPLSHLPLTGCLSTELQAAE 657
 DB 647 TNRRLSYLLTSLLC-----CFSSSLFIFGEPODWCRLRQPAFISFVLCISCLIVKTR 702
 QY 658 IFV--ESLPLS-----WADRLSGLRGPMAMLYVLLAMLYEVALCTIYLAAPPEVYT 710
 DB 703 VLVLEAKIPTSPHRKMWGLNQ-----FLVFLCTEQIYCALMTLWTPASSYRN 754
 QY 711 WHMLPTAL-VHCRTSRVWSFGLAHATNATLAFCLCTFLVRSOPGCVNRRAGLTFML 769
 DB 755 -HELEDEIIFITCHGSLMALGLIGYTCCLAALCFEFAFASRKLPENFNAKFTTFPSML 813

QY 770 AFTTWSEVELLANVQVLRPAVOMGALLCVGLLAAPHLPKCYLIMRQPGINTPE 827
 DB 814 IFFIWMISFIPAYASTGKFPVSAVEVIALIILAAAGGLACIFENNVYIILFKPSNTE 871

RESULT 11
 US-08-480-751-5
 : Sequence 5, Application US/08480751
 : Patent No. 5858684
 : GENERAL INFORMATION:
 : APPLICANT: Edward F. Nemeth
 : APPLICANT: Edward M. Brown
 : APPLICANT: Steven C. Hebert
 : APPLICANT: Forrest H. Fuller
 : APPLICANT: James E. Garrett, Jr.
 : TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 : NUMBER OF SEQUENCES: 20
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Lyon & Lyon
 : STREET: First Interstate World Center
 : STREET: Suite 4700
 : STREET: 633 West Fifth Street
 : CITY: Los Angeles
 : STATE: California
 : COUNTRY: USA
 : ZIP: 90071
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: FASTSEQ
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/480,751
 : FILING DATE: 7 June, 1995
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : PRIOR APPLICATION DATA: including application
 : PRIOR APPLICATION DATA: described below: 9
 : APPLICATION NUMBER: 08/353,784
 : FILING DATE: 9 December, 1994
 : APPLICATION NUMBER: PCT/US/94/12117
 : FILING DATE: 21 October, 1994
 : APPLICATION NUMBER: U.S. 08/292,827
 : FILING DATE: 23 August, 1994
 : APPLICATION NUMBER: U.S. 08/141,248
 : FILING DATE: 22 October, 1993
 : APPLICATION NUMBER: U.S. 08/009,389
 : FILING DATE: 23 February, 1993
 : APPLICATION NUMBER: U.S. 08/017,127
 : FILING DATE: 12 February, 1993
 : APPLICATION NUMBER: U.S. 07/934,161
 : FILING DATE: 21 August, 1992
 : APPLICATION NUMBER: U.S. 07/834,044
 : FILING DATE: 11 February, 1992
 : APPLICATION NUMBER: U.S. 07/749,451
 : FILING DATE: 23 August, 1991
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Hebert, Sheldon O.
 : REGISTRATION NUMBER: 38,179
 : REFERENCE/DOCKET NUMBER: 213/004
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (213) 489-1600
 : TELEFAX: (213) 955-0440
 : TELEX: 67-3510
 : INFORMATION FOR SEQ ID NO: 5:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1085 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-08-480-751-5

```

Query Match          23.4%; Score 1060; DB 2; Length 1085;
Best Local Similarity 30.0%; Pred. No. 1.2e-92;
Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

QY 6 VGLSLMALHFGTGAPLCSOOLRMKGDYVIGLFPPL--GEA--EBAIGRSRTRPSPVC 62
DB 10 ILAFSTWCTSAVGP-----DQRAOKKGDIIILGGLPIHFVAVKQDQDKS--RPSVVC 61
QY 63 TRFSSNGILMALAMKAAVEINNKSDLLPGLRLGDLPTDCSEPPVAMKPSIMFLA--KA 120
DB 62 IRYNFRGFRLQAMIFALFEINSSPALLPNTLGYRIPTCTVSKALATLISFVAQNKI 121
QY 121 GSRDLAAYCNYTOYOPRYLAVIGPSSSELAAMTGKFFSFELMPOYSYGASMEILSARETF 180
DB 122 DSIINDEFNCSEHIPSTIAVVGATGSGISTAVANILGLFTYPOYSYASSSRLSKNOF 181
QY 181 PSFERTVPSDRVQLTAAAEILQEFMNMVVAALGSDDEXRGOLISIFSALAAAGCICIAHE 240
DB 182 KSFLERTIPNDEHQATAMADIIEYFRNMVGTIAADDYGRPGIEKFRERAEERDIDERS 241
QY 241 GLVPLPRADDSRLGVQDVLHGVNOSSVQVLLFASVAAHMLFNTSISRLSPKWMAS 300
DB 242 ELI-----SOYSDEKIQOVVEYIOWSTAKVIVFSSGPDLEPLKEITVRNITGRIMLAS 297
QY 301 EAMLTSDIWMGLPGMAOMGTVLGFLORGAQLHEFPQYV-KTHALATDPAFCSAL----- 354
DB 298 EAMASSSLIAMEYFHVVGITIGFGLKAGQIQGFREFLOKVIHPRKSVNHGFAKEFEET 357
QY 355 -----GEREGL-----EDVVGQRPOCDCTILON 380
DB 356 NCHLOGANGPLVPDTFLNGHEGAGRLSNSPTAFRPLCTGEBNISVETPYMDYHLR- 416
QY 381 VSAGLNHQTSFYAAVYSVAQALNHTLOC-----NASGCPADDPKPMOLLEMMYN 432
DB 417 -----ISYVYLAIVSIALADITYCTICRGILFTNGS-CADIKKAYANQVLAHLKH 467
QY 433 LTFHVG-GILPLRDSGNDMEYDKLW--VMQGSVPLRHGVGRN--GSLTERL----- 483
DB 468 LMTFSNMGRQVTFDECGDLAGNYSIINMHLSPEDSGIV-FKEGVYNYVAKKGERLFIN 526
QY 484 -KIRNHTSNOKPVKSCSQCOEGOVRR-VKGFHSCCTCYVCEAGSYQNPDIDACTGC 541
DB 527 EKILMSGFSEVFPFSCSDCLAGTRKGIIEGPCCFECVBCPDEYSDETDASACDNC 586
QY 542 GODEMSPERSTCFRRRSRLAMGPAYVLLLLLSIALGLVLAAGLVVHHRDSPLYQA 601
DB 587 PDEFSNENHTCIAKEIFELTSMTEPGIALFLAVLGIETLAFVIGVITKRNPETIYVA 646
QY 602 SGGPLA--CFGLVLCGLVCL--SVLLFPGQSPARCLAQOPLSHLPLTCLSTLFLQAAE 657
DB 647 TNRELSTYLLFLSLC---CFSSSLFEGEPQDMTCRLQPAFGISFVLCISCIILVKNR 702
QY 658 IRY--ESELPLS-----WADRLSGCLRGPMWLVYLAAMEVALCTWYLVAPPEVTD 710
DB 703 VLLVFEAKIPTFSHRKMWGLNQ-----FLVFLCTFMQIVICALMLNAPSSSYRN 754
QY 711 WMLLPLEAL-VHCRTRSWVSGFLAHATNATLAFLELFGELVRSOPGCRNARGLFEAML 769
DB 755 -HELEDEIFITCHESLMALGFLIGYTCLLAICFFFAFKSRKRLDENENKAFITFSML 813
QY 770 AYTITWSEFPLLANVOVVLPRAYOMGALLCVLGLAFLHPRCYILMRQGLMTPPE 827
DB 814 IFFYIWISIPAYASTYKGFVSAVEVIALIASFGILACIFENKYYIILFKRSRMTE 871

```

```

; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,986
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/484,565
; FILING DATE: 7-June-1995
; APPLICATION NUMBER: 08/353,784
; FILING DATE: 9 December, 1994
; APPLICATION NUMBER: PCT/US/94/12117
; FILING DATE: 21 October, 1994
; APPLICATION NUMBER: U.S. 08/292,827
; FILING DATE: 23 August, 1994
; APPLICATION NUMBER: U.S. 08/141,248
; FILING DATE: 22 October, 1993
; APPLICATION NUMBER: U.S. 08/009,389
; FILING DATE: 23 February, 1993
; APPLICATION NUMBER: U.S. 08/017,127
; FILING DATE: 12 February, 1993
; APPLICATION NUMBER: U.S. 07/934,161
; FILING DATE: 21 August, 1992
; APPLICATION NUMBER: U.S. 07/834,044
; FILING DATE: 11 February, 1992
; APPLICATION NUMBER: U.S. 07/749,451
; FILING DATE: 23 August, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 213/006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELE: 67-3510
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1085 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-943-986-5

Query Match          23.4%; Score 1060; DB 2; Length 1085;
Best Local Similarity 30.0%; Pred. No. 1.2e-92;
Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

QY 6 VGLSLMALHFGTGAPLCSOOLRMKGDYVIGLFPPL--GEA--EBAIGRSRTRPSPVC 62
DB 10 ILAFSTWCTSAVGP-----DQRAOKKGDIIILGGLPIHFVAVKQDQDKS--RPSVVC 61
QY 63 TRFSSNGILMALAMKAAVEINNKSDLLPGLRLGDLPTDCSEPPVAMKPSIMFLA--KA 120
DB 62 IRYNFRGFRLQAMIFALFEINSSPALLPNTLGYRIPTCTVSKALATLISFVAQNKI 121
QY 121 GSRDLAAYCNYTOYOPRYLAVIGPSSSELAAMTGKFFSFELMPOYSYGASMEILSARETF 180
DB 122 DSIINDEFNCSEHIPSTIAVVGATGSGISTAVANILGLFTYPOYSYASSSRLSKNOF 181

```

```

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353,784
FILING DATE: 9 December, 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 8
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 209/069
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1085 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-784-5

Query Match 23.4% Score 1060; DB 3; Length 1085;
Best Local Similarity 30.0%; Pred. No. 1.2e-92;
Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

QY 6 VLGLMALLPRTGAPCLCQQLMKKGDYVLGGFPL--GEA--EAGLRSTRPSSPYC 62
DB 10 LLAFTWCTSAVGP-----DQRAKKKDDIILGGLFPIHFGVAVMDOLKS--RESVBC 61

QY 63 TRFSSNGLLMALAMKAAVEEINNKSDLLPGLRLGDTLFDTCSEPVYAMKPSIMFLA--KA 120
DB 62 IRYNFRGRFWQOAMFALTEEINSSPALLPNMTLGRIPDCTMTWSKALEATLSFAVAKMTI 121

QY 121 GSRIAAACNTOTQORVLAVIYGPSSSELAWTGKFFSFELMPQVSYGASMELLABRETF 180
DB 122 DSLNIDECNCSEHPSTTAVVAGATGSGISTRAVANLLGLFTYIPQVSYASSRLSNKNQF 181

QY 181 PSFRTVPSDVVOLTAALAEELLQEGFMNVAALGSDDEYGRGLSTFSALAAARGICIAHE 240
DB 182 KSLFTLIPNDHQATAMADIIETFRMNWVGIIAADDDVGRGIEKFRFEAEERDICIOPS 241

QY 241 GLVPLPRADSRIGKVDVYLHOVNOSVQVYLLFASVHAHAALFNYSISSRLSPVWVAS 300
DB 242 ELI-----SQYSDERKIQVVEYIONSTAKVIVFSSGPDLEPLLEIYVRNRTGRIWLAS 297

QY 301 EAWLTSDIVMGLPGMAQGVTLGFLQRAQOLHEEPQYV-KTHALATDPAPCSAL----- 354
DB 298 EAMASSSILAMPETFRHVVGITIGFLKAGQILGFRFELQKHVPRKSVNNGFAKEWEETF 357

QY 355 -----GEREQL-----EDVVGRCRQPCDCITLON 380

```



```

Db 358 NCHLOEAGAKGPLVDFTLHGHEGAGRLSNSPTARPLCTGEBNISVETPYMDYTHLR- 416
Qy 381 VSAGLNHQTSTVAAYVSAOALNHTLOC-----NASGCPADDPVKPQOLLENNYN 432
Db 417 -----ISYNYLVAYSIAHALODIYTCIGRGLFTNGS--CADIKKYEAMQVYKHLRH 467
Qy 433 LTFHVG-GIPLRFSSGNDMEYDKLW---VMGSSVPLHDVGRFN--GSLRTERL----- 483
Db 468 LNTSNMGEQVTFDECGDLAGNYSTIINMHLSPEDGSIV-FKEVGYNYVAKKGERLEIND 526
Qy 484 -KIRNHTSDNOKPVSRCSQCOEGQVRR-VKGFHSCCYDCVCEAGSYRONDDIACCTFC 541
Db 527 EKILMSGFREVFPNSCRDCLAGTRKGIIEBEPCCFCEVCPCPGEXSDETDASACDC 586
Qy 542 GODESPSRSTRCFRRRREFLAMEGPVALLLLSLALGLVLAALGLFVHHRDSPVQA 601
Db 567 PDDESNENHTSCIAKEIEFESWEFEIGALTLEFVAGIFLTAFLVGFIFRNPYKA 646
Qy 602 SGGPLA---CFGLVCLGLVCL-SVLLFPGQSPARCLAQOPLSHLPLTGCLSTLELQAAE 657
Db 647 TNRELSTYLLFLSLC-----CFSSSLFFIGEPQDWTCLRORAFGISFVLCISCLIVKTR 702
Qy 658 IFV--ESLPLS-----WADRISGCLRGPMAMLVLLMALVEVALCTWYLVAPPEVTD 710
Db 703 VLVFEAKIPTSFHKKWGLNLQ-----FLVFLCTFMQIVCAIMLNTAPSSRYN 754
Qy 711 WHMLPTEAL-VHCRTSRVSVFGLAHATNATLAFCELTFLVRSRPGCYNARGLTFEML 769
Db 755 -HELDELITFTCHGSLMALGFLGYCLLAICFFFAFRSRLKPENNENKFTFEFML 813
Qy 770 AYEITVSEVPLLANQVVLPRVOMGALLLCVLGILAFHLPRCLIMROGLTPE 827
Db 814 IFFIWIISFIPAVYSTGYKFSVAVEVIALIASFGLACIFENKYYIILFKRSNTIE 871

```

RESULT 14
US-08-484-7198-5
Sequence 5, Application US/084847198
Patent No. 6031003

GENERAL INFORMATION:

APPLICANT: Edward F. Nemeth, Edward M.
APPLICANT: Brown, Steven C. Hebert,
APPLICANT: Bradford C. Van Wagenen,
APPLICANT: Manuel F. Balandrin,
APPLICANT: Forrest H. Fuller, Eric G.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS Word
SOFTWARE: FASTSEQ for Windows Version 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,7198
FILING DATE: 7 June, 1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827

```

: FILING DATE: 23 August, 1994
: APPLICATION NUMBER: U.S. 08/141,248
: FILING DATE: 22 October, 1993
: APPLICATION NUMBER: U.S. 08/009,389
: FILING DATE: 23 February, 1993
: APPLICATION NUMBER: U.S. 08/017,127
: FILING DATE: 12 February, 1993
: APPLICATION NUMBER: U.S. 07/934,161
: FILING DATE: 21 August, 1992
: APPLICATION NUMBER: U.S. 07/834,044
: FILING DATE: 11 February, 1992
: APPLICATION NUMBER: U.S. 07/749,451
: FILING DATE: 23 August, 1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Douglas C. Murdock
: REGISTRATION NUMBER: 37,549
: REFERENCE/DOCKET NUMBER: 213/007
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: Telex: 67-3510
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1085 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-484-7198-5

```

Query Match 23.4%; Score 1060; DB 3; Length 1085;
Best Local Similarity 30.0%; Pred. No. 1.2e-92;
Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

```

Qy 6 VLGLSLALLHPGICAPLCSQQLRMKGDYVIGLFPPL--GEA-EEAGLRSTRSPSYC 62
Db 10 LIAFSTWCTSAVGP-----DQRAOKGDIILGGLFPIHFGVAVXDODLKS--RESEVYC 61
Qy 63 TRFSSNGLLMAMKMAVEEINNKSDLLPGLRLGDTDFDTGSEPVPVAMKPSLMFLA--KA 120
Db 62 IRYNRGRFRLQAMTFAIEEINSSPALLPNNLTGRIPTDCTVSKALAEATISFVAOKNT 121
Qy 121 GSRDIAAYCNVYQYQRYLVAVIGPSSSELAWTGKFFSEFLMPQVSYGASMELSARETF 180
Db 122 DSLNDEFRCNCEHPISTIAVVGATGSGISTAVANMLGLFYIPQVSYASSSLLENKNGF 181
Qy 181 PSFFRTVSDRYQLTAABELLOEFGMNVAALGSDDEYRGGLSTFSALAAAGCICIAHE 240
Db 182 KSFLRTIPDEHQATAMADIIIEYFWMNVGTITADDDYGRPGIEKFRERAEERDIDIFS 241
Qy 241 GLVPLPRADDSRLGKVODVLHQVNOSSVQVLLFASVNAHALFNYSSISRLSPVYVWS 300
Db 242 ELI-----SOYSPDEKIQOYVEVIONSTAKVIYVFSSGPLEPLIEIYARNITGRILAS 297
Qy 301 EAMLTSDLVMLGPGAOMGTVGLFQRGALHEFPQVY-KTHLALATPAFCSAL----- 354
Db 298 EAMASSSILAMPEYFHVVGCTIGFGIKAGQIGFREFLOKVPKRSVHNGFAKEWEETE 357
Qy 355 -----GEREGGL-----BEDVVGORCPDCCITLON 380
Db 358 NCHLOEAGAKGPLVDFTLHGHEGAGRLSNSPTARPLCTGEBNISVETPYMDYTHLR- 416
Qy 381 VSAGLNHQTSTVAAYVSAOALNHTLOC-----NASGCPADDPVKPQOLLENNYN 432
Db 417 -----ISYNYLVAYSIAHALODIYTCIGRGLFTNGS--CADIKKYEAMQVYKHLRH 467
Qy 433 LTFHVG-GIPLRFSSGNDMEYDKLW---VMGSSVPLHDVGRFN--GSLRTERL----- 483
Db 468 LNTSNMGEQVTFDECGDLAGNYSTIINMHLSPEDGSIV-FKEVGYNYVAKKGERLEIND 526
Qy 484 -KIRNHTSDNOKPVSRCSQCOEGQVRR-VKGFHSCCYDCVCEAGSYRONDDIACCTFC 541
Db 527 EKILMSGFREVFPNSCRDCLAGTRKGIIEBEPCCFCEVCPCPGEXSDETDASACDC 586

```

QY 542 GODEMSPERSTRCPRRSRFLANGEPVALLLLLSIALGLVLAALGLFVHHDRSPVQA 601
 Db 587 PDDFWSNENHTSCIAKEIEFISWTEPEGIALTEFAVIGIFLFAVIGIFENRNPYKA 646
 QY 602 SGGPLA---CGVLGLVCL-SVLEFPQGPSPARCLAQPLSHLPLTGLSTFLQAAE 657
 Db 647 TNRSLSTLLFSLC---CFSSSLFFGEGQDWTCLRPAGISFVLCISCLIVKTR 702
 QY 658 IFV--ESELPLS-----WADRLSGCLNGPMAVLLAMLEVALCWTYLAAPPEVTD 710
 Db 703 VLVLEFAKIPTRSHRKMGLNQ-----FLVFLCTFMQIVICAIWLTAPPSYRN 754
 QY 711 WHMLPTEAL-VHCRTRSWSEGLAHATNATLAFCLGTLFVRSOPGCYNRARGLTFAML 769
 Db 755 -HELEDELIIFTCHEGSLMALGFLIGYTCILAAICFFFAFKSRKLPENFNKAFITTESML 813
 QY 770 AYTFTWSEFVPLANNVOVLRPAVOMGALLCYGILAAFLPRCYLMLRPGCLNTPPE 827
 Db 814 IFTVWISFIPAVASTYGRKFSVAVEVIALIILASFGLLACIPFNKYIILFKPSRNTIE 871

RESULT 15

US-08-484-159-5

Sequence 5, Application US/08484159

Patent No. 6313146

GENERAL INFORMATION:

APPLICANT: Bradford C. Van Wagenen

APPLICANT: Manuel F. Balandrin

APPLICANT: Eric G. Del Mar

TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: First Interstate World Center

STREET: Suite 4700

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,159

FILING DATE: 7 June, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA: including application

Prior application data: described below: 9

APPLICATION NUMBER: 08/353,784

FILING DATE: 9 December, 1994

APPLICATION NUMBER: PCT/US/94/12117

FILING DATE: 21 October, 1994

APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994

APPLICATION NUMBER: U.S. 08/141,248

FILING DATE: 22 October, 1993

APPLICATION NUMBER: U.S. 08/009,389

FILING DATE: 23 February, 1993

APPLICATION NUMBER: U.S. 08/017,127

FILING DATE: 12 February, 1993

APPLICATION NUMBER: U.S. 07/934,161

FILING DATE: 21 August, 1992

APPLICATION NUMBER: U.S. 07/834,044

FILING DATE: 11 February, 1992

APPLICATION NUMBER: U.S. 07/749,451

FILING DATE: 23 August, 1991

ATTORNEY/AGENT INFORMATION:
 NAME: Heber, Sheldon O.
 REGISTRATION NUMBER: 38,179
 REFERENCE/DOCKET NUMBER: 214/101
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1085 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-484-159-5

Query Match 23.4%; Score 1060; DB 4; Length 1085;

Best Local Similarity 30.0%; Pred. No. 1,2e-92;

Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

QY 6 VLGSLMALHPGTGAPLCLISQOLRMKGDVYLGFLPL--GEA-EAAGLRSTRSPVC 62
 Db 10 ILAFSTWCTSAVGP-----DQRAQKGDITLGLFPHRGVAVKDDDLAS--RPESVEC 61
 QY 63 TRFSSNGILMALAMKMAVEEINNKSDLLPGRLGYDLPDCESEVYAMKPSLFLA--KA 120
 Db 62 IRYNRFGRFMQAMIFALIEEINSSPALLPNTLGRIFPDICNTYSKALENTLSVAQNKI 121
 QY 121 GSRDIAVCNTQYQYRVLAVIGPHSSELANVTKFSPFLMPOVYSGASHELLSARETF 180
 Db 122 DSLMDEFQNCSEHIPSTIAVVGATGSGISPAVANLGLFYIPQVYSASSRLLSNKNOF 181
 QY 181 PSFRTVPSDRVOLTAALAEILOEFQNMVVAALGSDDEYRGGLSIFALAAARGICIAHE 240
 Db 182 KSLRTIPNDHQATADIIETFRMNVGIIADDDYRGICIKFEAEAPRICIDFS 241
 QY 241 GLVPLPRADSRIGKVDVILQVNOSSVOYVLLFASVHAALFNYSISRLSPKVVAS 300
 Db 242 ELI-----SQYSDERKIQOVEVIONSTAKYIVVSSGPDLEPLIKEIYRRNITGRIVLAS 297
 QY 301 EAWLISDLVNGLPMAOMGVTLQGAQLHPRQV--KTHLALADDPACSL----- 354
 Db 298 EAWASSSLIAMPYEFHVAGGTIGFLKAGQIPGRERLQVHPKSVHNGPAAEFWEETP 357
 QY 355 -----GEREGGL-----EEDVYQORPOCCITLON 380
 Db 358 NCHLQEGAKGPLVDFTFLRHBEGGARLSPTAFRLCGEENISVEVPYNDYTLR- 416
 QY 381 VSAGLNHQTFSYAAVYVAQALHNTLQC-----NASGCPADDPKPMOLLERNYN 432
 Db 417 -----ISYNYLAVYSIAHLODIYTCIPGRGLFTNGS-CADIKKVAEMOVLKHLRH 467
 QY 433 LTFHVG-GLPLRPSSGNVMEYDLKLM--VMQGSVPRLDHGRFN-GSLRTERL----- 483
 Db 468 LNTFSMNGEYVTEDECDLGNSTIIMHLSPEDEGSLV-FKEGYIWNVYAKKGERLEIND 526
 QY 484 -KIRWHTSDNOKPYSRCSROCOEQVYR--VKGHSCCYDCEVDEAGSYRONPPDIATFC 541
 Db 527 EKILMSGFSREVPFNSRSCLAGTRKIGIEBGETCFCEVECPDGEYSDETASACDK 586
 QY 542 GODEMSPERSTRCPRRSRFLANGEPVALLLLLSIALGLVLAALGLFVHHDRSPVQA 601
 Db 587 PDDFWSNENHTSCIAKEIEFISWTEPEGIALTEFAVIGIFLFAVIGIFENRNPYKA 646
 QY 602 SGGPLA---CGVLGLVCL-SVLEFPQGPSPARCLAQPLSHLPLTGLSTFLQAAE 657
 Db 647 TNRSLSTLLFSLC---CFSSSLFFGEGQDWTCLRPAGISFVLCISCLIVKTR 702
 QY 658 IFV--ESELPLS-----WADRLSGCLNGPMAVLLAMLEVALCWTYLAAPPEVTD 710
 Db 703 VLVLEFAKIPTRSHRKMGLNQ-----FLVFLCTFMQIVICAIWLTAPPSYRN 754
 QY 711 WHMLPTEAL-VHCRTRSWSEGLAHATNATLAFCLGTLFVRSOPGCYNRARGLTFAML 769

Db 755 -HELEDEIIFITCHEGSLMALGFLIGYTCLLAICEFFAFKSRKLPENFNEAKFITESML 813
QY 770 AYEITWSEVPLANQVVLARPAVOMGAILLCVLIAGIILAFHLPKCYILMRQPLNTPPE 827
Db 814 IFFIWMISFIPAYASTYCKFVSAVEYAIILASFGILACIFENKVIYIILFKPSRNTTE 871

Search completed: May 19, 2003, 09:53:24
Job time : 27.6201 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:51:25 ; Search time 39.3211 Seconds
(without alignments)
2089.767 Million cell updates/sec

Title: US-09-927-315-15

Perfect score: 4524
Sequence: 1 MGPAVLGLSLALHPGTG.....GPGDAGQNDGNTGNCKRHE 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 362588 seqs, 96450795 residues

Total number of hits satisfying chosen parameters: 362588

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep: *
2: /cgn2_6/ptodata/2/pubpaa/PCR_NEW_PUB pep: *
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep: *
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep: *
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep: *
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep: *
7: /cgn2_6/ptodata/2/pubpaa/PCRUS_PUBCOMB pep: *
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep: *
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep: *
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep: *
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep: *
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep: *
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep: *
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4524	100.0	852	US-09-897-427A-6	Sequence 6, Appl1
2	4524	100.0	852	US-09-927-315-15	Sequence 15, Appl1
3	4524	100.0	852	US-10-035-045-4	Sequence 4, Appl1
4	4512	99.7	852	US-10-261-482-2	Sequence 2, Appl1
5	4512	99.7	852	US-10-287-837-14	Sequence 14, Appl1
6	4512	99.7	852	US-09-786-338A-14	Sequence 4, Appl1
7	4500	99.5	850	US-09-799-629-4	Sequence 14, Appl1
8	3255	71.9	858	US-09-799-629-14	Sequence 14, Appl1
9	3255	71.9	858	US-09-927-315-25	Sequence 25, Appl1
10	3255	71.9	858	US-10-035-045-14	Sequence 14, Appl1
11	3229	71.4	858	US-09-927-315-23	Sequence 23, Appl1
12	3227.5	71.3	858	US-09-927-315-18	Sequence 18, Appl1
13	3221.5	71.2	858	US-09-927-315-20	Sequence 20, Appl1
14	1769	39.1	1138	US-10-261-482-4	Sequence 4, Appl1
15	1208.5	26.7	840	US-09-361-652-1	Sequence 1, Appl1
16	1208.5	26.7	840	US-09-927-315-1	Sequence 2, Appl1
17	1188.5	26.3	842	US-09-361-652-2	Sequence 2, Appl1
18	1188.5	26.3	842	US-09-927-315-2	Sequence 2, Appl1
19	1153	25.3	841	US-09-897-427A-2	Sequence 2, Appl1

20	1153	25.5	841	US-09-799-629-17	Sequence 17, Appl1
21	1153	25.5	841	US-10-035-045-17	Sequence 17, Appl1
22	1152	25.5	841	US-09-819-946-2	Sequence 2, Appl1
23	1082.5	23.9	763	US-09-819-946-4	Sequence 4, Appl1
24	1068.5	23.6	1078	US-09-727-205-2	Sequence 2, Appl1
25	1068.5	23.6	1078	US-10-002-854-2	Sequence 2, Appl1
26	1067.5	23.6	1078	US-10-125-792-28	Sequence 28, Appl1
27	1067.5	23.6	1078	US-10-125-778-28	Sequence 28, Appl1
28	1065	23.5	941	US-10-125-792-8	Sequence 8, Appl1
29	1065	23.5	941	US-10-125-778-8	Sequence 8, Appl1
30	1064	23.5	941	US-10-125-792-10	Sequence 10, Appl1
31	1064	23.5	941	US-10-125-778-10	Sequence 10, Appl1
32	1056	23.3	1027	US-10-125-792-2	Sequence 2, Appl1
33	1056	23.3	1027	US-10-125-778-2	Sequence 2, Appl1
34	1053	23.3	777	US-09-361-652-3	Sequence 3, Appl1
35	1053	23.3	777	US-09-927-315-3	Sequence 3, Appl1
36	1048.5	23.2	850	US-10-125-792-12	Sequence 12, Appl1
37	1048.5	23.2	850	US-10-125-778-12	Sequence 12, Appl1
38	1027.5	22.7	926	US-09-816-685-2	Sequence 2, Appl1
39	1014.5	22.4	839	US-09-897-427A-4	Sequence 4, Appl1
40	1014.5	22.4	839	US-10-035-045-21	Sequence 21, Appl1
41	1011	22.3	838	US-09-927-315-9	Sequence 9, Appl1
42	1008.5	22.3	828	US-09-816-685-4	Sequence 4, Appl1
43	1002.5	22.2	843	US-10-096-144-1	Sequence 1, Appl1
44	1002.5	22.2	843	US-09-927-315-7	Sequence 7, Appl1
45	1002.5	22.2	843	US-10-124-598-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-897-427A-6
; Sequence 6, Application US/09897427A
; Patent No. US20020160424A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIADONG
; APPLICANT: STALEWSKI, LENA
; APPLICANT: XU HONG
; APPLICANT: EHEVERRI, FERNANDO
; TITLE OF INVENTION: T1R HETERO-OLIGOMERIC TASTE RECEPTORS
; FILE REFERENCE: 078003-0282558
; CURRENT APPLICATION NUMBER: US/09/897, 427A
; CURRENT FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-427A-6

Query Match	100.0%	Score 4524;	DB 9;	Length 852;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 852; Conservative	0;	Mismatches	0;	Gaps 0;
DB	1	MGPAVLGLSLALHPGTGAPICLSQOLRMKGDVVLGFLGRAEVAGRSRPPSP	60	
QY	1	MGPAVLGLSLALHPGTGAPICLSQOLRMKGDVVLGFLGRAEVAGRSRPPSP	60	
DB	61	VCTRRSSNGLMALMKNAVEIRNKSDLLPLRLGYDFTDCSPVYAMKPSLFLAKA	120	
QY	61	VCTRRSSNGLMALMKNAVEIRNKSDLLPLRLGYDFTDCSPVYAMKPSLFLAKA	120	
DB	121	GSRDIAAICNTYQYOPRYAVIGPSSSEIATWYGFSSFFLMPQVSYGASMLLSARETF	180	
QY	121	GSRDIAAICNTYQYOPRYAVIGPSSSEIATWYGFSSFFLMPQVSYGASMLLSARETF	180	
DB	181	PSFFFTVSDRQVLAALAEELQEFGMNVAALGSDDEYRGSLTFSLAARGLCIAHE	240	
QY	181	PSFFFTVSDRQVLAALAEELQEFGMNVAALGSDDEYRGSLTFSLAARGLCIAHE	240	
DB	181	PSFFFTVSDRQVLAALAEELQEFGMNVAALGSDDEYRGSLTFSLAARGLCIAHE	240	

```

QY 241 GLVPLPRADSRDLGKVDVYLHQVNOSSVQVYVLLFASVHAHALNYSISSLSRPRVWAS 300
D 241 GLVPLPRADSRDLGKVDVYLHQVNOSSVQVYVLLFASVHAHALNYSISSLSRPRVWAS 300
QY 301 EAWLTSIDLVMGLPGMAOMGTVLGFLQKGAOLHEFPQYKTHLATATDPAFCALSAGEROG 360
D 301 EAWLTSIDLVMGLPGMAOMGTVLGFLQKGAOLHEFPQYKTHLATATDPAFCALSAGEROG 360
QY 361 LEEDVVGRCPOCCITLQNVASAGLNHHQTFESYAAYVSAVALHNTLQCNASGCPADP 420
D 361 LEEDVVGRCPOCCITLQNVASAGLNHHQTFESYAAYVSAVALHNTLQCNASGCPADP 420
QY 421 VKPQOLLENNMTNLFHFHVGGLPLRFDSGNDMEYDLKLMWQGSVPRLDHVGRENGSLRT 480
D 421 VKPQOLLENNMTNLFHFHVGGLPLRFDSGNDMEYDLKLMWQGSVPRLDHVGRENGSLRT 480
QY 481 ERLKIRMTSDNOKPVSRCSQCOEGOVRRYKGFHSCCYDCVDEAGSYRONPDIACTF 540
D 481 ERLKIRMTSDNOKPVSRCSQCOEGOVRRYKGFHSCCYDCVDEAGSYRONPDIACTF 540
QY 541 CGODEWSPERSTRCFRRRSRFLANGEPAYVLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
D 541 CGODEWSPERSTRCFRRRSRFLANGEPAYVLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
QY 601 ASGPPLACFGIVCLGIVCLSVLTFPGQSPARCIAQOPLSHLPLTGCLSTFLQAAETV 660
D 601 ASGPPLACFGIVCLGIVCLSVLTFPGQSPARCIAQOPLSHLPLTGCLSTFLQAAETV 660
QY 661 ESELPLSWADRLSGCLNGPMAVLVLLAMLYEVALCTWYLAFFPEVYTDHMLPTETLV 720
D 661 ESELPLSWADRLSGCLNGPMAVLVLLAMLYEVALCTWYLAFFPEVYTDHMLPTETLV 720
QY 721 HCRTRSWVSFGLAHATNATLAFICFLGTFIVRSOPGCGYNRARGLTFAMLAFTWVSFP 780
D 721 HCRTRSWVSFGLAHATNATLAFICFLGTFIVRSOPGCGYNRARGLTFAMLAFTWVSFP 780
QY 781 LLAVQVYLRAVOMGALLLCVGLTIAAFHLPRCYLLMRQGLNTPEFFLGCGGDAQO 840
D 781 LLAVQVYLRAVOMGALLLCVGLTIAAFHLPRCYLLMRQGLNTPEFFLGCGGDAQO 840
QY 841 NDGNTGNQKHE 852
D 841 NDGNTGNQKHE 852

```

RESULT 2

```

US-09-927-315-15
; Sequence 15, Application US/09927315
; Publication No. US20030040045A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nicholas J.P.
; APPLICANT: Nelson, Greg
; APPLICANT: Hoon, Mark A.
; APPLICANT: Chandrasekar, Jayaram
; APPLICANT: Zhang, Yifeng
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
; FILE REFERENCE: 02307E-12011005
; CURRENT APPLICATION NUMBER: US/09/927,315
; PRIOR APPLICATION NUMBER: US 60/302,898
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 15
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; OTHER INFORMATION: human T1R3 sweet taste receptor
US-09-927-315-15

```

```

Query Match 100.0%; Score 4524; DB 9; Length 852;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MLGPAVLGLSLMALHGTGAPLCLSOOLRMKGDYVGLGLEPLGEAEAGLRSTRSPSP 60
D 1 MLGPAVLGLSLMALHGTGAPLCLSOOLRMKGDYVGLGLEPLGEAEAGLRSTRSPSP 60
QY 61 VCTRSSNGLLMALAMKMAVEINNKSDLLPLRLGIDLEFTYCEPVAAMPSTLATAKA 120
D 61 VCTRSSNGLLMALAMKMAVEINNKSDLLPLRLGIDLEFTYCEPVAAMPSTLATAKA 120
QY 121 GSRDIAYCANTYQOPRLAYIGPHSSSLAMVTKFSPFLMPQVSGASMEILLSARETF 180
D 121 GSRDIAYCANTYQOPRLAYIGPHSSSLAMVTKFSPFLMPQVSGASMEILLSARETF 180
QY 181 PSFRTVPDSRVOQLTAAAEILLQEEFGMMVVAALGSDDEYGRGLSIFSLAAARGICIAHE 240
D 181 PSFRTVPDSRVOQLTAAAEILLQEEFGMMVVAALGSDDEYGRGLSIFSLAAARGICIAHE 240
QY 241 GLVPLPRADSRDLGKVDVYLHQVNOSSVQVYVLLFASVHAHALNYSISSLSRPRVWAS 300
D 241 GLVPLPRADSRDLGKVDVYLHQVNOSSVQVYVLLFASVHAHALNYSISSLSRPRVWAS 300
QY 301 EAWLTSIDLVMGLPGMAOMGTVLGFLQKGAOLHEFPQYKTHLATATDPAFCALSAGEROG 360
D 301 EAWLTSIDLVMGLPGMAOMGTVLGFLQKGAOLHEFPQYKTHLATATDPAFCALSAGEROG 360
QY 361 LEEDVVGRCPOCCITLQNVASAGLNHHQTFESYAAYVSAVALHNTLQCNASGCPADP 420
D 361 LEEDVVGRCPOCCITLQNVASAGLNHHQTFESYAAYVSAVALHNTLQCNASGCPADP 420
QY 421 VKPQOLLENNMTNLFHFHVGGLPLRFDSGNDMEYDLKLMWQGSVPRLDHVGRENGSLRT 480
D 421 VKPQOLLENNMTNLFHFHVGGLPLRFDSGNDMEYDLKLMWQGSVPRLDHVGRENGSLRT 480
QY 481 ERLKIRMTSDNOKPVSRCSQCOEGOVRRYKGFHSCCYDCVDEAGSYRONPDIACTF 540
D 481 ERLKIRMTSDNOKPVSRCSQCOEGOVRRYKGFHSCCYDCVDEAGSYRONPDIACTF 540
QY 541 CGODEWSPERSTRCFRRRSRFLANGEPAYVLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
D 541 CGODEWSPERSTRCFRRRSRFLANGEPAYVLLLLSLALGLVLAALGLFVHHRDSPLVQ 600
QY 601 ASGPPLACFGIVCLGIVCLSVLTFPGQSPARCIAQOPLSHLPLTGCLSTFLQAAETV 660
D 601 ASGPPLACFGIVCLGIVCLSVLTFPGQSPARCIAQOPLSHLPLTGCLSTFLQAAETV 660
QY 661 ESELPLSWADRLSGCLNGPMAVLVLLAMLYEVALCTWYLAFFPEVYTDHMLPTETLV 720
D 661 ESELPLSWADRLSGCLNGPMAVLVLLAMLYEVALCTWYLAFFPEVYTDHMLPTETLV 720
QY 721 HCRTRSWVSFGLAHATNATLAFICFLGTFIVRSOPGCGYNRARGLTFAMLAFTWVSFP 780
D 721 HCRTRSWVSFGLAHATNATLAFICFLGTFIVRSOPGCGYNRARGLTFAMLAFTWVSFP 780
QY 781 LLAVQVYLRAVOMGALLLCVGLTIAAFHLPRCYLLMRQGLNTPEFFLGCGGDAQO 840
D 781 LLAVQVYLRAVOMGALLLCVGLTIAAFHLPRCYLLMRQGLNTPEFFLGCGGDAQO 840
QY 841 NDGNTGNQKHE 852
D 841 NDGNTGNQKHE 852

```

RESULT 3

```

US-10-035-045-4
; Sequence 4, Application US/10035045
; Publication No. US20030054448A1
; GENERAL INFORMATION:

```

```

; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIAODONG
; APPLICANT: STASZEWSKI, LENA
; APPLICANT: O'CONNELL, SHAWN
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: T1R TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003-0280681
; CURRENT APPLICATION NUMBER: US/10/035,045
; CURRENT FILING DATE: 2002-01-03
; PRIOR APPLICATION NUMBER: 60/259,227
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: 60/284,547
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-035-045-4

```

```

Query Match          100.0%; Score 4524; DB 9; Length 852;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MCGPAVLGSLWALLHPGTGAPCLISQOLRMKGDYVIGLFLPGEAEAGLSRTRPSSP 60
DB 1 MCGPAVLGSLWALLHPGTGAPCLISQOLRMKGDYVIGLFLPGEAEAGLSRTRPSSP 60
QY 61 VCTRESSNGLLWALMKMAVEEINNKSDLLPGIRLGYDLDFPCSEPVVAMKPSLMFLARA 120
DB 61 VCTRESSNGLLWALMKMAVEEINNKSDLLPGIRLGYDLDFPCSEPVVAMKPSLMFLARA 120
QY 121 GSRDIAAYCNYTOYOPRVLAVIGPHSSELMATYTKFFSFLMPQVSYGASMEILSARETF 180
DB 121 GSRDIAAYCNYTOYOPRVLAVIGPHSSELMATYTKFFSFLMPQVSYGASMEILSARETF 180
QY 181 PSFFRTVPDRVOLTAAAEELQEFQWNVVAAALGSDDEYGRGISTIFSALAAAGCICIAHE 240
DB 181 PSFFRTVPDRVOLTAAAEELQEFQWNVVAAALGSDDEYGRGISTIFSALAAAGCICIAHE 240
QY 241 GLVPLPRADDSRLGKQVDLHOVNOSSVOYVLLFASVHAHAHLFNTSISRSPKXWVAS 300
DB 241 GLVPLPRADDSRLGKQVDLHOVNOSSVOYVLLFASVHAHAHLFNTSISRSPKXWVAS 300
QY 301 EAMLISDLVWGLPGMAOMGTVLGFLORGAQLHEFPQYVTHALATDPAFCALGEREG 360
DB 301 EAMLISDLVWGLPGMAOMGTVLGFLORGAQLHEFPQYVTHALATDPAFCALGEREG 360
QY 361 LEEDVVGORCPQDCITLTONVSAGLNHQTFSVYAAVSVAAALNHTLQCNASGCPADP 420
DB 361 LEEDVVGORCPQDCITLTONVSAGLNHQTFSVYAAVSVAAALNHTLQCNASGCPADP 420
QY 421 VAPWOLLEMMYMLTFHVGLPLRFDSGAVDMEYDKLWMOGSPVRLHDVGRFNGSLRT 480
DB 421 VAPWOLLEMMYMLTFHVGLPLRFDSGAVDMEYDKLWMOGSPVRLHDVGRFNGSLRT 480
QY 481 EBLKTRMHNSDNOKPVSRCROCGOVRVYKGFHSCCDVDCGEGSYRONDDIACGF 540
DB 481 EBLKTRMHNSDNOKPVSRCROCGOVRVYKGFHSCCDVDCGEGSYRONDDIACGF 540
QY 541 CGODEMSPSTRCFRRRSRFLAMGEPAYLILLLILSLALGLVLAALGLFVHHRSPLVQ 600
DB 541 CGODEMSPSTRCFRRRSRFLAMGEPAYLILLLILSLALGLVLAALGLFVHHRSPLVQ 600
QY 601 ASGGPLACGLVCLGVCLSVLLFPQSPARCLAOQPLSHPLRGCLSTFLQAAEIV 660
DB 601 ASGGPLACGLVCLGVCLSVLLFPQSPARCLAOQPLSHPLRGCLSTFLQAAEIV 660
QY 661 ESELPLSNADRLSGCLRGFAMLVYLLALVEALCTWTLVAPPPVYVDWMLPREALV 720
DB 661 ESELPLSNADRLSGCLRGFAMLVYLLALVEALCTWTLVAPPPVYVDWMLPREALV 720

```

```

QY 721 HCRTRSWSEGLAHATNATLAFCLFGLFVRSOPCYNRARGTLFAMLAFTTWSEVP 780
DB 721 HCRTRSWSEGLAHATNATLAFCLFGLFVRSOPCYNRARGTLFAMLAFTTWSEVP 780
QY 781 LLANVOVVLPAVOMGALLVGLIAAEHLPRCYLLMRQPLNPEFLGGGPGDAQO 840
DB 781 LLANVOVVLPAVOMGALLVGLIAAEHLPRCYLLMRQPLNPEFLGGGPGDAQO 840
QY 841 NNGNTGNCKRHE 852
DB 841 NNGNTGNCKRHE 852

```

RESULT 4

```

US-10-261-482-2
; Sequence 2, Application US/10261482
; Publication No. US20030036089A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: CLO00869CON
; CURRENT APPLICATION NUMBER: US/10/261,482
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 09/684,393
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: 60/172,600
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 852
; TYPE: PRT
; ORGANISM: HUMAN
US-10-261-482-2

```

```

Query Match          99.7%; Score 4512; DB 9; Length 852;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 851; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MCGPAVLGSLWALLHPGTGAPCLISQOLRMKGDYVIGLFLPGEAEAGLSRTRPSSP 60
DB 1 MCGPAVLGSLWALLHPGTGAPCLISQOLRMKGDYVIGLFLPGEAEAGLSRTRPSSP 60
QY 61 VCTRESSNGLLWALMKMAVEEINNKSDLLPGIRLGYDLDFPCSEPVVAMKPSLMFLARA 120
DB 61 VCTRESSNGLLWALMKMAVEEINNKSDLLPGIRLGYDLDFPCSEPVVAMKPSLMFLARA 120
QY 121 GSRDIAAYCNYTOYOPRVLAVIGPHSSELMATYTKFFSFLMPQVSYGASMEILSARETF 180
DB 121 GSRDIAAYCNYTOYOPRVLAVIGPHSSELMATYTKFFSFLMPQVSYGASMEILSARETF 180
QY 181 PSFFRTVPDRVOLTAAAEELQEFQWNVVAAALGSDDEYGRGISTIFSALAAAGCICIAHE 240
DB 181 PSFFRTVPDRVOLTAAAEELQEFQWNVVAAALGSDDEYGRGISTIFSALAAAGCICIAHE 240
QY 241 GLVPLPRADDSRLGKQVDLHOVNOSSVOYVLLFASVHAHAHLFNTSISRSPKXWVAS 300
DB 241 GLVPLPRADDSRLGKQVDLHOVNOSSVOYVLLFASVHAHAHLFNTSISRSPKXWVAS 300
QY 301 EAMLISDLVWGLPGMAOMGTVLGFLORGAQLHEFPQYVTHALATDPAFCALGEREG 360
DB 301 EAMLISDLVWGLPGMAOMGTVLGFLORGAQLHEFPQYVTHALATDPAFCALGEREG 360
QY 361 LEEDVVGORCPQDCITLTONVSAGLNHQTFSVYAAVSVAAALNHTLQCNASGCPADP 420
DB 361 LEEDVVGORCPQDCITLTONVSAGLNHQTFSVYAAVSVAAALNHTLQCNASGCPADP 420
QY 421 VAPWOLLEMMYMLTFHVGLPLRFDSGAVDMEYDKLWMOGSPVRLHDVGRFNGSLRT 480
DB 421 VAPWOLLEMMYMLTFHVGLPLRFDSGAVDMEYDKLWMOGSPVRLHDVGRFNGSLRT 480

```

QY 481 ERLKIRMTSDNOKPVSRCSROCOEGVRYKGFHSCCYDCVDEAGSYRONPDIACF 540
 DB 481 ERLKIRMTSDNOKPVSRCSROCOEGVRYKGFHSCCYDCVDEAGSYRONPDIACF 540
 QY 541 CGODEMSPERSTRCFRRRSRLANGEPAYVLLLLSLALGLVLAALGLFVHHHDSPLYQ 600
 DB 541 CGODEMSPERSTRCFRRRSRLANGEPAYVLLLLSLALGLVLAALGLFVHHHDSPLYQ 600
 QY 601 ASGGPLACFGIYVCLGYLSTLFPGOPSAPARCLAOQPLSHPLTGCLSTFLQAAETV 660
 DB 601 ASGGPLACFGIYVCLGYLSTLFPGOPSAPARCLAOQPLSHPLTGCLSTFLQAAETV 660
 QY 661 ESELPLSWADRLSGCLRGPMAMLVLLAMLEVALCTWYLAAPPEVVTDMHMLPTALV 720
 DB 661 ESELPLSWADRLSGCLRGPMAMLVLLAMLEVALCTWYLAAPPEVVTDMHMLPTALV 720
 QY 721 HCRTRSWVSFGIAHATNTATLFLCFLGTFIVRSOPGGRYNRARGLTFAMLAFTWVSFVP 780
 DB 721 HCRTRSWVSFGIAHATNTATLFLCFLGTFIVRSOPGGRYNRARGLTFAMLAFTWVSFVP 780
 QY 781 LLANVQVYLRPAVOMGALLLCVLTGLIAAFHLPCYLLMRQGLNTPPEFLGGGPDAGQ 840
 DB 781 LLANVQVYLRPAVOMGALLLCVLTGLIAAFHLPCYLLMRQGLNTPPEFLGGGPDAGQ 840
 QY 841 NDGNTNGNGKHE 852
 DB 841 NDGNTNGNGKHE 852

RESULT 5

US-10-282-837-14
 : Sequence 14, Application US/10282837
 : Publication No. US20030082738A1
 : GENERAL INFORMATION:
 : APPLICANT: Millennium Pharmaceuticals, Inc.
 : TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
 : TITLE OF INVENTION: 52872, NOVEL G PROTEIN-COUPLED RECEPTORS AND USES THEREFOR
 : FILE REFERENCE: 10448-020001
 : CURRENT APPLICATION NUMBER: US/10/282, 837
 : PRIOR FILING DATE: 2002-10-29
 : PRIOR APPLICATION NUMBER: US/09/796, 338
 : PRIOR FILING DATE: 2001-02-28
 : PRIOR APPLICATION NUMBER: US 60/186, 059
 : NUMBER OF SEQ ID NOS: 26
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 14
 : LENGTH: 852
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-10-282-837-14

Query Match 99.7%; Score 4512; DB 9; Length 852;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 851; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MCGPAVIGLSIMALLHPGTGAPLCLSQOLRMKGDTVLGGLPPLGEAEAGLRSTRPSP 60
 DB 1 MCGPAVIGLSIMALLHPGTGAPLCLSQOLRMKGDTVLGGLPPLGEAEAGLRSTRPSP 60
 QY 61 VCTRSSNGILMALAMKAIVEINKKSDLPGLRGYDLPTCSPVYAMPSLMEFLAKA 120
 DB 61 VCTRSSNGILMALAMKAIVEINKKSDLPGLRGYDLPTCSPVYAMPSLMEFLAKA 120
 QY 121 GSRDIAACNTYQOPRYLATVIGPHSSSLAMVTGKFFSFLMPQVSYGASMEILSARETF 180
 DB 121 GSRDIAACNTYQOPRYLATVIGPHSSSLAMVTGKFFSFLMPQVSYGASMEILSARETF 180
 QY 181 PSFRTVPSDRVQLTAAELLQERFGWNVVALGSDDEXGROGLSTFSALAAAGICIAHE 240
 DB 181 PSFRTVPSDRVQLTAAELLQERFGWNVVALGSDDEXGROGLSTFSALAAAGICIAHE 240
 QY 241 GLVPLPRADSDRLGKVQVYLAHQVNOSSVOYVLLPASVAHAALPNYSISSRLSPKVVAS 300

DB 241 GLVPLPRADSDRLGKVQVYLAHQVNOSSVOYVLLPASVAHAALPNYSISSRLSPKVVAS 300
 QY 301 EAMLTSDIYMGILPGMAOMGTVLGFLONGAOLHEEPQYKTHLATATPAFCALGEREG 360
 DB 301 EAMLTSDIYMGILPGMAOMGTVLGFLONGAOLHEEPQYKTHLATATPAFCALGEREG 360
 QY 361 LEEDVQQRPCQDCITLQVNSAGLNHHQTFESVYAAYVAQALHNTLQCNASGCPADP 420
 DB 361 LEEDVQQRPCQDCITLQVNSAGLNHHQTFESVYAAYVAQALHNTLQCNASGCPADP 420
 QY 421 VKPMQLENNMTNLTFFHYGGLPLRFDSSGNVMEIDLKIMWQGSVPRLDHGRNGSLRT 480
 DB 421 VKPMQLENNMTNLTFFHYGGLPLRFDSSGNVMEIDLKIMWQGSVPRLDHGRNGSLRT 480
 QY 481 ERLKIRMTSDNOKPVSRCSROCOEGVRYKGFHSCCYDCVDEAGSYRONPDIACF 540
 DB 481 ERLKIRMTSDNOKPVSRCSROCOEGVRYKGFHSCCYDCVDEAGSYRONPDIACF 540
 QY 541 CGODEMSPERSTRCFRRRSRLANGEPAYVLLLLSLALGLVLAALGLFVHHHDSPLYQ 600
 DB 541 CGODEMSPERSTRCFRRRSRLANGEPAYVLLLLSLALGLVLAALGLFVHHHDSPLYQ 600
 QY 601 ASGGPLACFGIYVCLGYLSTLFPGOPSAPARCLAOQPLSHPLTGCLSTFLQAAETV 660
 DB 601 ASGGPLACFGIYVCLGYLSTLFPGOPSAPARCLAOQPLSHPLTGCLSTFLQAAETV 660
 QY 661 ESELPLSWADRLSGCLRGPMAMLVLLAMLEVALCTWYLAAPPEVVTDMHMLPTALV 720
 DB 661 ESELPLSWADRLSGCLRGPMAMLVLLAMLEVALCTWYLAAPPEVVTDMHMLPTALV 720
 QY 721 HCRTRSWVSFGIAHATNTATLFLCFLGTFIVRSOPGGRYNRARGLTFAMLAFTWVSFVP 780
 DB 721 HCRTRSWVSFGIAHATNTATLFLCFLGTFIVRSOPGGRYNRARGLTFAMLAFTWVSFVP 780
 QY 781 LLANVQVYLRPAVOMGALLLCVLTGLIAAFHLPCYLLMRQGLNTPPEFLGGGPDAGQ 840
 DB 781 LLANVQVYLRPAVOMGALLLCVLTGLIAAFHLPCYLLMRQGLNTPPEFLGGGPDAGQ 840
 QY 841 NDGNTNGNGKHE 852
 DB 841 NDGNTNGNGKHE 852

RESULT 6

US-09-796-338A-14
 : Sequence 14, Application US/09796338A
 : Patent No. US20020061522A1
 : GENERAL INFORMATION:
 : APPLICANT: Millennium Pharmaceuticals, Inc.
 : TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
 : TITLE OF INVENTION: 52872, NOVEL G PROTEIN-COUPLED RECEPTORS AND USES THEREFOR
 : FILE REFERENCE: 10448-020001
 : CURRENT APPLICATION NUMBER: US/09/796, 338A
 : PRIOR FILING DATE: 2001-02-28
 : PRIOR APPLICATION NUMBER: US 60/186, 059
 : NUMBER OF SEQ ID NOS: 26
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 14
 : LENGTH: 852
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-09-796-338A-14

Query Match 99.7%; Score 4512; DB 10; Length 852;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 851; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MCGPAVIGLSIMALLHPGTGAPLCLSQOLRMKGDTVLGGLPPLGEAEAGLRSTRPSP 60
 DB 1 MCGPAVIGLSIMALLHPGTGAPLCLSQOLRMKGDTVLGGLPPLGEAEAGLRSTRPSP 60

		99.5%	Score 4500:	DB 9:	Length 850:	
Query Match	Best Local Similarity	99.6%	Pred. No. 0:			
Matches 849:	Conservative	0:	Mismatches 1:	Indels 2:	Gaps 1:	
US-09-799-629-4						
?	ORGANISM: Homo sapiens					
?	TYPE: PRT					
?	LENGTH: 850					
?	SEQ ID NO: 4					
?	SOFTWARE: Patentln Ver. 2.1					
?	NUMBER OF SEQ ID NOS: 20					
?	PRIOR FILING DATE: 2001-01-03					
?	PRIOR APPLICATION NUMBER: 60/259,227					
?	PRIOR FILING DATE: 2000-06-23					
?	PRIOR APPLICATION NUMBER: 60/214,213					
?	PRIOR FILING DATE: 2000-06-23					
?	PRIOR APPLICATION NUMBER: 60/226,448					
?	PRIOR FILING DATE: 2000-08-17					
?	PRIOR APPLICATION NUMBER: 60/259,227					
?	PRIOR FILING DATE: 2001-01-03					
?	NUMBER OF SEQ ID NOS: 20					
?	SOFTWARE: Patentln Ver. 2.1					
?	SEQ ID NO: 4					
?	LENGTH: 850					
?	TYPE: PRT					
?	ORGANISM: Homo sapiens					
?	US-09-799-629-4					
Query Match	Best Local Similarity	99.5%	Score 4500:	DB 9:	Length 850:	
Matches 849:	Conservative	0:	Mismatches 1:	Indels 2:	Gaps 1:	
DB	1	MGPPVVLGSLWALLHPGTGAPLCLISQOLRMGDIVLGGFLPGLGEAEAGLSRTRPSSP	60			
DB	1	MGPPVVLGSLWALLHPGTGAPLCLISQOLRMGDIVLGGFLPGLGEAEAGLSRTRPSSP	60			
QY	61	VCTRRSSNGLLMALMKAAVEEINKNSDLLPGLRIGYDLFTPCSEPPYAMKSLFLAKA	120			
DB	61	VCTRRSSNGLLMALMKAAVEEINKNSDLLPGLRIGYDLFTPCSEPPYAMKSLFLAKA	120			
QY	121	GSRDIAAFCNTYQYQPRVLAIVGPHSSSELAMTYGKFFSEFLAPQVSYGASMEILLSARETF	180			
DB	121	GSRDIAAFCNTYQYQPRVLAIVGPHSSSELAMTYGKFFSEFLAPQVSYGASMEILLSARETF	180			
QY	121	GSRDIAAFCNTYQYQPRVLAIVGPHSSSELAMTYGKFFSEFLAPQVSYGASMEILLSARETF	180			
DB	121	GSRDIAAFCNTYQYQPRVLAIVGPHSSSELAMTYGKFFSEFLAPQVSYGASMEILLSARETF	180			
QY	181	PSFFRTVSDRYQVLTAAAEILLOEFGNNVYALGSDDEYTRQGLSTFSALAAARGICIAHE	240			
DB	179	PSFFRTVSDRYQVLTAAAEILLOEFGNNVYALGSDDEYTRQGLSTFSALAAARGICIAHE	238			
QY	241	GLVPLPRADSRIRGKQVDLVHQNOSYQVYLLFASVHAHLFNTSISRSLSPKVMVAS	300			
DB	239	GLVPLPRADSRIRGKQVDLVHQNOSYQVYLLFASVHAHLFNTSISRSLSPKVMVAS	298			
QY	301	EAMVLTSDLVYMLPGMAOKGTVYGLQRGACQLHEFPQYVYKTHALATDPAFCALSAGEREG	360			
DB	299	EAMVLTSDLVYMLPGMAOKGTVYGLQRGACQLHEFPQYVYKTHALATDPAFCALSAGEREG	358			
QY	361	LEEDVVGOCPCDDCTTLOVNSAGLNHOFESVYAAYVSVAOALNHTLOCNSSGCAADP	420			
DB	359	LEEDVVGOCPCDDCTTLOVNSAGLNHOFESVYAAYVSVAOALNHTLOCNSSGCAADP	418			
QY	421	VKPMOLLEMYMLTFHVHVGILPLRFDSGCVNDMEYDKLKWVMOGSVRLHDVGRFNGSLRT	480			
DB	419	VKPMOLLEMYMLTFHVHVGILPLRFDSGCVNDMEYDKLKWVMOGSVRLHDVGRFNGSLRT	478			
QY	481	EKLKTRWHTSDNQKQVYSRCSRQOCQGVARVYGVHSCCYDVCDEAGSTRONPDIACTF	540			
DB	479	EKLKTRWHTSDNQKQVYSRCSRQOCQGVARVYGVHSCCYDVCDEAGSTRONPDIACTF	538			
QY	541	CGODMSPSRSTRRCRRRRRFLAMGEPVALLLLLLLSTALGLVLAALGFVHHRRDSPVQ	600			
DB	539	CGODMSPSRSTRRCRRRRRFLAMGEPVALLLLLLLSTALGLVLAALGFVHHRRDSPVQ	598			
QY	601	ASGGPLACGVLGVLCLSVLLPFGQPSPARCLAQDPLSHLPLTGCLSTFLQAAEIFV	660			
DB	599	ASGGPLACGVLGVLCLSVLLPFGQPSPARCLAQDPLSHLPLTGCLSTFLQAAEIFV	658			
QY	661	ESELPLSNADRLSGCLRGWMAVLVYLLMLVYALCTWLVAFPPPVYVDMHMLPPEALY	720			
DB	659	ESELPLSNADRLSGCLRGWMAVLVYLLMLVYALCTWLVAFPPPVYVDMHMLPPEALY	718			
QY	721	HCRTSWVSFGIAHATNNTLAFLCELTGTFVLSQPGCYKARAGLTFEAMLAVERITYWSEVP	780			
DB	719	HCRTSWVSFGIAHATNNTLAFLCELTGTFVLSQPGCYKARAGLTFEAMLAVERITYWSEVP	778			

QY 781 L1ANVOVLRPAVOMGALLLCVIGIILAAFLHPRCYLLMRPGMLTPEFLGGPGDAOQ 840
 Db 779 L1ANVOVLRPAVOMGALLLCVIGIILAAFLHPRCYLLMRPGMLTPEFLGGPGDAOQ 838
 QY 841 NDGNTGNOKHE 852
 Db 839 NDGNTGNOKHE 850

RESULT 8

US-09-799-629-14
 ; Sequence 14, Application US/09799629
 ; Publication No. US20030008344A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ADLER, JON ELLIOT
 ; APPLICANT: ZOZULYA, SERGEY
 ; APPLICANT: LI, XIADONG
 ; APPLICANT: O'CONNELL, SHAWN
 ; APPLICANT: STASEWSKI, LENA
 ; TITLE OF INVENTION: T1R TASTE RECEPTORS AND GENES ENCODING SAME
 ; FILE REFERENCE: 078003/0277870/RXT
 ; CURRENT APPLICATION NUMBER: US/09/799, 629
 ; PRIOR FILING DATE: 2001-03-07
 ; PRIOR APPLICATION NUMBER: 60/187,546
 ; PRIOR FILING DATE: 2000-03-07
 ; PRIOR APPLICATION NUMBER: 60/195,536
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: 60/209,840
 ; PRIOR FILING DATE: 2000-06-06
 ; PRIOR APPLICATION NUMBER: 60/214,213
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: 60/226,448
 ; PRIOR FILING DATE: 2000-08-17
 ; PRIOR APPLICATION NUMBER: 60/259,227
 ; PRIOR FILING DATE: 2001-01-03
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 858
 ; TYPE: PRT
 ; ORGANISM: Rattus sp.
 US-09-799-629-14

Query Match 71.9%; Score 3255; DB 9; Length 858;
 Best Local Similarity 72.9%; Pred. No. 2.8e-254;
 Matches 623; Conservative 74; Mismatches 144; Indels 14; Gaps 4;

QY 1 MLCRAVIGISIMALLHGTGAPICLSOOLRMKGDYVIGLFLPGAEEAGLRSTRSPSP 60
 Db 1 MGLAIIIGLSIAAFLLEGMSSICLSOOFKAGDYILGGLEPLGTTEATLNORTOPNGI 60
 QY 61 VCTFRSSNGLLMALAMKAAVEINNKSDLPGLRGYDLFTPCSEPPVYAMPSLMFLAKA 120
 Db 61 LCTFRSPGLFLAAMKAAVEINNGSLLPLGLRGYDLFTPCSEPPVYAMPSLMFLAKA 120
 QY 121 GSRDIAAYCANTYOYOPRVLAVIIGHSESLAVNTGKFSFFLMPOVSYGASMSLARETF 180
 Db 121 GSOSIAAYCANTYOYOPRVLAVIIGHSESLAITGKFSFFLMPOVSYGASMSLARETF 180
 QY 181 PSFRTPVSDRVOJTAABEILLEGCGWVVAALGSDDEYRGGLSIFSLAARIGICIAHE 240
 Db 181 PSFRTPVSDRVOJTAABEILLEGCGWVVAALGSDDEYRGGLSIFSLAARIGICIAHE 240
 QY 241 GLVPLPRADBSRLGKVDYLVHVNQSSVOYVLLFASVHAALHNSYSSLSKRVWVAS 300
 Db 241 GLVPOHDTSGQLGKVDYLVHVNQSSVOYVLLFASARAVYSLSTSLHLSKRVWVAS 300
 QY 301 EAMLTSDLVMLPGMAOMGTVLGFLRGGAOLHEPPQYVYKTHALATDPACFALGERBOG 360
 Db 301 ESMLTSDLVMLPNIARVGLGFLRGGALLPERSHYVETRLAALADPTFCASL-KAELD 359
 QY 361 LEEVYQRCRCQCCTITLQNSAGLN-----HQTFSYAAVYSVAQALHNTLQCN 412

Db 360 LEEVYQRCRCQCCTITLQNSAGLN-----HQTFSYAAVYSVAQALHNTLQCNV 419
 QY 413 SCGPADOPVAPWOLLENNYILTFHVGLPLRFDSGNGVDEYDLKLMVMOGSVPLRDVYG 472
 Db 420 SHCHTSEPVPQWOLENNYMKSPFARDLTQTFDAKGSVDMEYDLKLMVMOGSVPLRVHTVG 479
 QY 473 RFNGSLTERIKIRMTSDNOKPVSRCSROCGVRRVKGFSHCCYDCVCEAGSTRON 532
 Db 480 TFNGTLOHSSKMYW--PGNOVPVSCSROCKDQGVRRVKGFSHCCYDCVCEAGSTRON 537
 QY 533 PDDIACFPCCODEMSPRSTRRCFRBRFLAMGPATVLLILLSTALGLVLAALGLFVH 592
 Db 538 PDDTCTPCCKDQWSPKSTCTCPRRPRFLAMGPATVLLILLSTALGLVLAALGLFVH 597
 QY 593 HRDSPLOVAGSGPLACGLVLCGLVCLVYLLFPQSPARCLAOQPSHPLTGCSTLFL 652
 Db 598 YWDSPLVQASGSLFCGGLICLGLFCLSVLLFPRPSPASACLAQPPAHPLPITGCLSTLF 657
 QY 653 LQAAEFVESELPLSMADRLSGCLRGPMVLVLLAMLEVALCTWYLVAPPEVYDWH 712
 Db 658 LQAAEFVESELPLSMANWLCSTYLRGPMVLVLLATVLEALCAWYLVAPPEVYDWH 717
 QY 713 MLPTFALVHCRTBSVSPGLAHATNATLPLCIGTFLVYSQRCYRARGITPAMLAYF 772
 Db 718 VLPTEVLEHCRMRWSVSLGVHTINAVLAFLCFLGTFLVOSQPRRYNRARGITPAMLAYF 777
 QY 773 ITWVSFPLLANVOVLRPAVOMGALLLCVIGIILAAFLHPRCYLLMRPGMLTPEFLGG 832
 Db 778 ITWVSFPLLANVOVLRPAVOMGALLLCVIGIILAAFLHPRCYLLMRPGMLTPEFLGG 837
 QY 833 GPGDAGQNDGNTGN 847
 Db 838 SPKEA---SDGNSGS 849

RESULT 9

US-09-927-315-25
 ; Sequence 25, Application US/09927315
 ; Publication No. US20030040045A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zuker, Charles S.
 ; APPLICANT: Ryba, Nicholas J.P.
 ; APPLICANT: Nelson, Greg
 ; APPLICANT: Hoon, Mark A.
 ; APPLICANT: Chandrasekar, Jayaram
 ; APPLICANT: Zhang, Yifeng
 ; APPLICANT: The Regents of the University of California
 ; APPLICANT: The Government of the United States of America
 ; APPLICANT: as represented by the Secretary of the
 ; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
 ; FILE REFERENCE: 02307E-120110US
 ; CURRENT APPLICATION NUMBER: US/09/927,315
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: US 60/302,898
 ; PRIOR FILING DATE: 2001-07-03
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 25
 ; LENGTH: 858
 ; TYPE: PRT
 ; ORGANISM: Rattus sp.
 ; FEATURE:
 ; OTHER INFORMATION: rat T1R3 sweet taste receptor
 US-09-927-315-25

Query Match 71.9%; Score 3255; DB 9; Length 858;
 Best Local Similarity 72.9%; Pred. No. 2.8e-254;
 Matches 623; Conservative 74; Mismatches 144; Indels 14; Gaps 4;
 QY 1 MLCRAVIGISIMALLHGTGAPICLSOOLRMKGDYVIGLFLPGAEEAGLRSTRSPSP 60
 Db 1 MGLAIIIGLSIAAFLLEGMSSICLSOOFKAGDYILGGLEPLGTTEATLNORTOPNGI 60

Db 1 MGLAIIIGLSIAFLAELGNGSSSLCSQOFKAGDYLIGLFLPGTEATLNRQPNGI 60
QY 61 VCTRSSNGLMALAMKMAVEEINKNSDILLPGLRLGYDLPDPCSEPVYAMKSLMFLA 120
Db 61 ICTRSSPLGLFLAMAKMAVEEINKNSALLPGLRLGYDLPDPCSEPVYAMKSLMFLA 120
QY 121 GSRDIAAYCNTYOYOPRYLAIVIGPHSSELAMTGFKEFSEFLMPOVYSASMDLSRETF 180
Db 121 GSQSLAAYCNTYOYOPRYLAIVIGPHSSELALITGKFEFSEFLMPOVYSASMDLSRETF 180
QY 181 PSFRTVPDRVOLTAALAEELQEFGNMVAALGSDDEXGROGLSTFSALAARGICIAHE 240
Db 181 PSFRTVPDRVOLTAALAEELQEFGNMVAALGSDDEXGROGLSTFSALAARGICIAHE 240
QY 241 GLVPLPRADDSRLKGYODVLHOVNOSSVOYVLLFASVHAHAHLFNYSISSRLSPKMWAS 300
Db 241 GLVPHDHTSGQOLGKYVDVLROVNOSSKYOVVLLFASARAVYLSFYSIILHDSIPKMWAS 300
QY 301 EAMLTSDLVNGLPMAOMGTVLGFLQGAQLHEFPQYVTHALATDPAFCSALGEREG 360
Db 301 EAMLTSDLVNGLPMAOMGTVLGFLQGAQLHEFPQYVTHALATDPAFCSALGEREG 360
QY 361 LEEDVVGORCPQCCITLONVSNAGLN-----HHQFTSYAAYVSAQAALHNTLQCN 412
Db 361 LEEDVVGORCPQCCITLONVSNAGLN-----HHQFTSYAAYVSAQAALHNTLQCN 412
QY 413 SCCPADPVPKPMQLENNYNLTFEHVGLPLRDESSGNVMEVDLKMVWQSPVPLHTVG 472
Db 413 SCCPADPVPKPMQLENNYNLTFEHVGLPLRDESSGNVMEVDLKMVWQSPVPLHTVG 472
QY 473 RENGSLRTELRKIRWHTSDNOKPVSRCSROCEGOVRRYKGFHSCCYDCVDEAGSYRON 532
Db 473 RENGSLRTELRKIRWHTSDNOKPVSRCSROCEGOVRRYKGFHSCCYDCVDEAGSYRON 532
QY 532 TENGTLQLOHSMY--PENOVPVSGCSROCKDGOVRRYKGFHSCCYDCVDEAGSYRON 537
Db 532 TENGTLQLOHSMY--PENOVPVSGCSROCKDGOVRRYKGFHSCCYDCVDEAGSYRON 537
QY 533 PDDIACTEGQDSEPSERSTRCFRRSRFLANGEPVAVLLLLLSLALGLVLAALGLFVH 592
Db 533 PDDIACTEGQDSEPSERSTRCFRRSRFLANGEPVAVLLLLLSLALGLVLAALGLFVH 592
QY 538 HRDSPVQASGGPLACFLGVLCLVCLVLPFGOPSPARCIAQOPLSLPLTGLSTLF 652
Db 538 HRDSPVQASGGPLACFLGVLCLVCLVLPFGOPSPARCIAQOPLSLPLTGLSTLF 652
QY 653 LOAAEIFESELPLSMANMLCSYLRGPAMLVLLATLVEALCALMYLMAFPEVYTDNQ 712
Db 653 LOAAEIFESELPLSMANMLCSYLRGPAMLVLLATLVEALCALMYLMAFPEVYTDNQ 712
QY 713 MLPTALVHCRTSRVSWFGLAHATNATLAFCLGTEFLVRSOPGCYNRARGITFAMLAYF 772
Db 713 MLPTALVHCRTSRVSWFGLAHATNATLAFCLGTEFLVRSOPGCYNRARGITFAMLAYF 772
QY 773 ITWVSFVPLANVOYVLAIVAVOMGAILLCVGLIILAFHLPRCYLMLROGLMTPPEFLG 832
Db 773 ITWVSFVPLANVOYVLAIVAVOMGAILLCVGLIILAFHLPRCYLMLROGLMTPPEFLG 832
QY 833 GPDAQGQNDGNTGN 847
Db 833 SPKRA---SDNGSGS 849

RESULT 10
US-10-035-045-14
; Sequence 14, Application US/10035045
; Publication No. US20030054448A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIAODONG
; APPLICANT: STASZEWSKI, LENA
; APPLICANT: O'CONNELL, SHAWN
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: TIR TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003-0280681
; CURRENT APPLICATION NUMBER: US/10/035, 045
; CURRENT FILING DATE: 2002-01-03

QY 1 MGLPAVILGSLIALMHPGAPLCSLQOIRMKGDVYLGGLFPLGEAEELGNSRTPSSP 60
Db 1 MGLPAVILGSLIALMHPGAPLCSLQOIRMKGDVYLGGLFPLGEAEELGNSRTPSSP 60
QY 61 VCTRSSNGLMALAMKMAVEEINKNSDILLPGLRLGYDLPDPCSEPVYAMKSLMFLA 120
Db 61 ICTRSSPLGLFLAMAKMAVEEINKNSALLPGLRLGYDLPDPCSEPVYAMKSLMFLA 120
QY 121 GSRDIAAYCNTYOYOPRYLAIVIGPHSSELAMTGFKEFSEFLMPOVYSASMDLSRETF 180
Db 121 GSQSLAAYCNTYOYOPRYLAIVIGPHSSELALITGKFEFSEFLMPOVYSASMDLSRETF 180
QY 181 PSFRTVPDRVOLTAALAEELQEFGNMVAALGSDDEXGROGLSTFSALAARGICIAHE 240
Db 181 PSFRTVPDRVOLTAALAEELQEFGNMVAALGSDDEXGROGLSTFSALAARGICIAHE 240
QY 241 GLVPLPRADDSRLKGYODVLHOVNOSSVOYVLLFASVHAHAHLFNYSISSRLSPKMWAS 300
Db 241 GLVPHDHTSGQOLGKYVDVLROVNOSSKYOVVLLFASARAVYLSFYSIILHDSIPKMWAS 300
QY 301 EAMLTSDLVNGLPMAOMGTVLGFLQGAQLHEFPQYVTHALATDPAFCSALGEREG 360
Db 301 EAMLTSDLVNGLPMAOMGTVLGFLQGAQLHEFPQYVTHALATDPAFCSALGEREG 360
QY 361 LEEDVVGORCPQCCITLONVSNAGLN-----HHQFTSYAAYVSAQAALHNTLQCN 412
Db 361 LEEDVVGORCPQCCITLONVSNAGLN-----HHQFTSYAAYVSAQAALHNTLQCN 412
QY 413 SCCPADPVPKPMQLENNYNLTFEHVGLPLRDESSGNVMEVDLKMVWQSPVPLHTVG 472
Db 413 SCCPADPVPKPMQLENNYNLTFEHVGLPLRDESSGNVMEVDLKMVWQSPVPLHTVG 472
QY 473 RENGSLRTELRKIRWHTSDNOKPVSRCSROCEGOVRRYKGFHSCCYDCVDEAGSYRON 532
Db 473 RENGSLRTELRKIRWHTSDNOKPVSRCSROCEGOVRRYKGFHSCCYDCVDEAGSYRON 532
QY 532 TENGTLQLOHSMY--PENOVPVSGCSROCKDGOVRRYKGFHSCCYDCVDEAGSYRON 537
Db 532 TENGTLQLOHSMY--PENOVPVSGCSROCKDGOVRRYKGFHSCCYDCVDEAGSYRON 537
QY 533 PDDIACTEGQDSEPSERSTRCFRRSRFLANGEPVAVLLLLLSLALGLVLAALGLFVH 592
Db 533 PDDIACTEGQDSEPSERSTRCFRRSRFLANGEPVAVLLLLLSLALGLVLAALGLFVH 592
QY 538 HRDSPVQASGGPLACFLGVLCLVCLVLPFGOPSPARCIAQOPLSLPLTGLSTLF 652
Db 538 HRDSPVQASGGPLACFLGVLCLVCLVLPFGOPSPARCIAQOPLSLPLTGLSTLF 652
QY 653 LOAAEIFESELPLSMANMLCSYLRGPAMLVLLATLVEALCALMYLMAFPEVYTDNQ 712
Db 653 LOAAEIFESELPLSMANMLCSYLRGPAMLVLLATLVEALCALMYLMAFPEVYTDNQ 712
QY 713 MLPTALVHCRTSRVSWFGLAHATNATLAFCLGTEFLVRSOPGCYNRARGITFAMLAYF 772
Db 713 MLPTALVHCRTSRVSWFGLAHATNATLAFCLGTEFLVRSOPGCYNRARGITFAMLAYF 772
QY 773 ITWVSFVPLANVOYVLAIVAVOMGAILLCVGLIILAFHLPRCYLMLROGLMTPPEFLG 832
Db 773 ITWVSFVPLANVOYVLAIVAVOMGAILLCVGLIILAFHLPRCYLMLROGLMTPPEFLG 832
QY 833 GPDAQGQNDGNTGN 847

Query Match 71.9% Score 3255; DB 9; Length 858;
Best Local Similarity 72.9%; Pred. No. 2,8e-254;
Matches 623; Conservative 74; Mismatches 144; Indels 14; Gaps 4;
; PRIOR APPLICATION NUMBER: 60/259, 227
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: 60/284, 547
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 858
; TYPE: PRT
; ORGANISM: Rattus sp.
; US-10-035-045-14

RESULT 14
US-10-261-482-4
; Sequence 4, Application US/10261482
; Publication No. US20030036089A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CLO008693CON
; CURRENT APPLICATION NUMBER: US/10/261,482
; PRIOR APPLICATION NUMBER: 09/684,393
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: 60/172,600
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1138
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-261-482-4

Query Match 39.1%; Score 1769; DB 9; Length 1138;
Best Local Similarity 37.1%; Pred. No. 4,6e-134;
Matches 427; Conservative 103; Mismatches 254; Indels 368; Gaps 16;

QY 31 MKGCVYLGLEPPL-GEAEAGLRSTRSPVCTR---FSSNGILMALMAKVAEINN- 85
DB 2 LRGFLAGLFLSLGDC---LQYRHRPLVYSCRPDSFNHGHILHQAIFYEEINNS 57
QY 86 KSDLLPGLRLGYDLFTCSSEPVAMKPSLMFLARAGSRDIAANCYQYQPRVLAVTGP 145
DB 58 QSDLLPGLRLGYDLFTCSSEPVAMKPSLMFLARAGSRDIAANCYQYQPRVLAVTGP 117
QY 146 S----- 146
DB 118 SSLPLGLTDCSELLAGRIVAGPSSALLPNTILGYELYDVCSANVATRLVLAQPR 177
QY 147 -----SELAMTGRKFFSEFLMPQVSYGASMLLSARETFP 181
DB 178 RHIEIQRDHRHSSKVVAFIPDNOSELAMTGRKFFSEFLMPQVSYGASMLLSARETFP 237
QY 182 SEFTVPSDRQLTAAAELOEFGH----- 206
DB 238 SEFTVPSDRQLTAAAELOEFGHMAFLMPVSYASLSAPPSFTVPSDRQLTQFGWSTD 297
QY 207 -----N 207
DB 298 HAVTTAALLGPFLLMPVSYEASSVYLSAKKFPFLRTVPSDRHQUEVYVQQLQSGMGN 357
QY 208 WVAALGSDDEXGROGLSTFSALAAARGICIAHEGLVPLPRADSRRLGVQDVLHQVN--- 264
DB 358 WVAALGSDDEXGROGLSTFSALAAARGICIAHEGLVPLPRADSRRLGVQDVLHQVNQNG 417
QY 265 ----- 264
DB 418 SYGLIARGICAVPRDOOSWISLIGSYGYGLGVOALEELAVPRGICVAFKDIPEFSA 477
QY 265 -----OSSVOVLLFASVHAHALFNYSISSRLSPRYWASEAMLSIDL 308
DB 478 RVGDPFRMOSKMOHLAQQSSVOVLLFASVHAHALFNYSISSRLSPRYWASEAMLSIDL 537
QY 309 VMGLPGMAQMGTVLGFQRCQAOLHEF----- 334
DB 538 VMGLPGMAQMGTVLGF-----VFHAFLKVMVASFMSGTVLGATTVYVVSNNHLARVF 594
QY 335 -----POYVKTALATPAPFCSALGEROGLHEE 363
DB 595 FRSVYLANLGKVMVASSEDMAISTYITSVIGIGITVGLVAVQOQVPGKKEEESYVR 654
QY 364 DV-----VGQRPCQDCITLQNVS--AGLNHHQFVSVAAYVSAQALHNT 407

DB 655 AVTAAPSACPEGSWCSNQLCRECHFTTRNMPITGAFSMAAAYVEAYVAHGHQL 714
QY 408 LQNASGCPADDPVPMQMLLEMTNLTFFHVGGLRLRDSGSSNVMEYDKLMMQGSVPR 467
DB 715 LGCTSEIC-SRGVYPMQLQIITKYNFLHENTVARDNDGDTGYDIIAMDNGPEWT 773
QY 468 LHDVGRFNGS---LATERLKIWMHTSDNQKPEVSRCSROCGQVRYRGKGFHSCCYDCVD 524
DB 774 FEIGSASLSPVHLDINKTKIQHGNKNQVSVCTDCLAGHRRVYVGHSCFECVPC 833
QY 525 EAGSYRONPDDIACFEGQDSEWSPERSTRCRRSRLTANGEPVALLILSLALGLV 584
DB 834 EAGFTLNMSELHIOQPCGTEWMAKESTCPRTVEFLAMEPISTVLIAANTLILLLV 893
QY 585 AALGFVHHRDSPVQASGGLPACGLVCLVCLVSLVLPFGQSPARLAAQPLSHLPL 644
DB 894 GTAGLFAMHFTPYVRASGGLCFMLGSLVAGSCSYSPFGPTVACILRQLFSLG 953
QY 645 TGCLSTFLQAAELFV---ESLPL---SWADRLSGCLRQPMAMVLLAMLEVALCT 697
DB 954 AIFLSCLTIRSFOLVITFKFSTKVPTFRTWA-----QNHGAGLRFVYSSVHLILCL 1006
QY 698 WYLAFPPEVYTDHMLPTENLVHCRTRSWYSEGLATNTATLFLCFGLFVYRSQPC 757
DB 1007 TWLVMMTPRPTRQYRPHVILDECTEVSVGLFAAFTHTLISITFVCSYLKELPEN 1066
QY 758 YNRARGLEFAMLAFTVTSVFPPLANVQVVLRAVOMGALLICVLGILAFHLPRLCYL 817
DB 1067 YNEAKCTFSLILNFVSMIAFTMASTYQGSYLRAVAVYLAGLTLTSGFSGLPKCTVI 1126
QY 818 MRQPLNTPPEF 829
DB 1127 LCRLPRLNTEHF 1138

RESULT 15
US-09-361-652-1
; Sequence 1, Application US/09361652
; Publication No. US2003003630A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Lindemeyer, Juergen
; APPLICANT: Ryda, Nick
; APPLICANT: Hoon, Mark
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
; FILE REFERENCE: 02307E-088610US
; CURRENT APPLICATION NUMBER: US/09/361,652
; CURRENT FILING DATE: 1999-07-27
; EARLIER APPLICATION NUMBER: US 60/094,465
; EARLIER FILING DATE: 1998-07-28
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentl Ver. 2.1
; SEQ ID NO 1
; LENGTH: 840
; TYPE: PRT
; ORGANISM: Rattus sp.
; OTHER INFORMATION: rat G-protein coupled receptor B3 (GPCR-B3)
US-09-361-652-1

Query Match 26.7%; Score 1208.5; DB 9; Length 840;
Best Local Similarity 33.8%; Pred. No. 5.7e-89;
Matches 291; Conservative 151; Mismatches 339; Indels 79; Gaps 18;

QY 12 WALHPGTGAPLCLSDQL-----RKKGCVYLGLEPPL-GEAEAGLR 53
DB 4 WA-----AHLLSLDLYVCMWAFSCQRTSSPQSLPGLAGLFLSLGDC---LQY 52
QY 54 RTRPSSPVCTR---FSSNGILMALMAKVAEINNKSDDLPLGRLGYDLFTCSSEPVAM 110

```
Db      53  RHRPLVTSQDRDSEFNGHGYHLFOAMRFVVEEINNSSALLPNTLTGLEYDYCVSES-ANY 111
QY      111  KPSLMEFLAKAGSRDIAYCNTOYOPRYLAVNGPHSSELAMTGKFFSFFLMPQVSYGAS 170
Db      112  YATLRVLAIOGPRHIEIOKDLRNHSSKVVAFIGPDMDHAYTTALLGPFLMPLVSEAS 171
QY      171  MELLASARETFPSFEFTVPDSRVOLTAABELLOEFGWNVVAALGSDDEYGRGCLSFSA 230
Db      172  SYVLSAKRKFPSPFLRVPSDRQVEVMVOLLOSFGWVWISLIGSYDYGOLGVALLEELA 231
QY      231  AARGICIAHEGLVPLERADDSRLG--KVODVLHQNQSSVQVLLFASVAAHALEFNYSI 288
Db      232  VPRGICVAFKDIYPF-----SARVGDPRMQSMQHIAQARTTVVYVFSNRHLARVFFRSYV 287
QY      289  SSRISPKNVAVASEAMTSLDVNGLPQMAQMGTYGLQGAQ--LHEFPQ--YKTHLAL 344
Db      288  LANLTGKVVAVASEDMAISTYITSVTGIQIGTYLGAVOQROVPGIKEFEESYVRAVTA 347
QY      345  AT---DPAFCSALGEREQGLEEDVVGQRCPOCDITQNVS--AGLNHHQTFSVYAAYS 399
Db      348  PSACPEBGSWCS-----TNQLCRECHFTTRNMPTLGAFSMAATRYEAVYA 394
QY      400  VAOALHNTIOCNAGCPCPADPYKPMQLENNMYNLFFHVGGPLRFDSGNDMEYDKLM 459
Db      395  VAHGLHQLLGCTSEIC-SRGPYPMQILOQIYKVNFLHENTYVAFDDNGDTLGYDIIAM 453
QY      460  VMQGSVPRHLDVGRFNGS---LTERLKTIRWHTSDNQKPVSCSRQCQEGQVRRKRGHS 516
Db      454  DNGPBMWTFEIIIGSASLSLVHLDINKTKIOHGNKNQVPSVCTDCLAGHHRVVGSHH 513
QY      517  CCYDCVDCBAGSYRONPDIDACTFCGODEMSPERSTRCFRRSRFLAWEPAVLLLLLL 576
Db      514  CCFECVPCBAGTFLMNSSELHIOPCGTEEMAKESTTCEPRIVERFLAMHEPISIVLIAM 573
QY      577  STALGLVLAALGLFVHHNDSPLVOASGGLACGLVCLGLVCLISVLEFPQSPARCIAQ 636
Db      574  TLILLLVGTAGLFAWHFHTPVVRSAGRLCFLMLGSLVAGSCSFYSFGEPTVPACLIR 633
QY      637  OPLSHLPLTGCLSTLFLQAAEIFV---ESELPL---SWADRLSGCLRPMAMIVLLAM 689
Db      634  QPLFSIGFAIFISCLTIRSFOLVITFKSTKVPTFYRTWA-----QNHGAGLEVYBS 686
QY      690  LVEVALCTWYLVAFPEPVVTDHMLPTALVHCRTSRWVSFGIAHATNATLAFLCLGTF 749
Db      687  FVHLICTLFWLVMTPRPRFREQRPFLVILBCTEVNSVGFLLAFTHNILLSISTFVCSY 746
QY      750  LYRSQPGCYNRARGLTFAMLAFTITWVSFVPLIANQVYLRAVOMGALLLCVLTIAAF 809
Db      747  LKRELPEENTNEAKCVTFSLINFSWIAFTWASITQGSYLPVAVNLAGLTTLISGGSFSGY 806
QY      810  HLPKCYILMROPGLNTPPEF 829
Db      807  FLPKCYIILCRPELNNTHEF 826
```

Search completed: May 19, 2003, 10:04:45
Job time : 43.8231 secs

THIS PAGE BLANK (USPTO)